



HAL
open science

Variabilités émotionnelle et neurophysiologique dans la perception olfactive : étude par l'intelligence artificielle

Maelle Moranges

► To cite this version:

Maelle Moranges. Variabilités émotionnelle et neurophysiologique dans la perception olfactive : étude par l'intelligence artificielle. Neurosciences. Université Claude Bernard - Lyon I, 2023. Français. NNT : 2023LYO10009 . tel-04432891

HAL Id: tel-04432891

<https://theses.hal.science/tel-04432891v1>

Submitted on 1 Feb 2024

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.



**THESE de DOCTORAT
DE L'UNIVERSITE CLAUDE BERNARD LYON 1**

**Ecole Doctorale 476
(ED Neurosciences et Cognition)**

Spécialité de doctorat : Neurosciences

Soutenue publiquement le 18/01/2023, par :
Maëlle Moranges

**Variabilités émotionnelle
et neurophysiologique
dans la perception olfactive :
étude par l'intelligence artificielle**

Devant le jury composé de :

Mohand-Saïd Hacid
Peggy Cellier
Denis Pierron
Caroline Huart
Céline Rouveïrol
Moustafa Bensafi
Marc Plantevit

Professeur des Universités, Univ. Lyon 1
Maîtresse de conférences HDR, INSA Rennes
Chargé de recherche, CNRS Toulouse
Professeure, Univ. Catholique de Louvain
Professeure des Universités, Univ. Paris 13
Directeur de Recherche, CNRS Lyon
Professeur, EPITA

Président
Rapporteuse
Rapporteur
Examinatrice
Examinatrice
Directeur de thèse
Co-directeur de thèse

Université Claude Bernard - LYON 1

Président de l'Université	M. Frédéric FLEURY
Président du Conseil Académique	M. Hamda BEN HADID
Vice-Président du Conseil d'Administration	M. Didier REVEL
Vice-Président du Conseil des Etudes et de la Vie Universitaire	M. Philippe CHEVALLIER
Vice-Président de la Commission de Recherche	M. Petru MIRONESCU
Directeur Général des Services	M. Pierre ROLLAND

COMPOSANTES SANTE

Département de Formation et Centre de Recherche en Biologie Humaine	Directrice : Mme Anne-Marie SCHOTT
Faculté d'Odontologie	Doyenne : Mme Dominique SEUX
Faculté de Médecine et Maïeutique Lyon Sud - Charles Mérieux	Doyenne : Mme Carole BURILLON
Faculté de Médecine Lyon-Est	Doyen : M. Gilles RODE
Institut des Sciences et Techniques de la Réadaptation	Directeur : M. Xavier PERROT
Institut des Sciences Pharmaceutiques et Biologiques	Directrice : Mme Christine VINCIGUERRA

COMPOSANTES & DEPARTEMENTS DE SCIENCES & TECHNOLOGIE

Département Génie Electrique et des Procédés	Directrice : Mme Rosaria FERRIGNO
Département Informatique	Directeur : M. Behzad SHARIAT
Département Mécanique	Directeur M. Marc BUFFAT
Ecole Supérieure de Chimie, Physique, Electronique	Directeur : Gérard PIGNAULT
Institut de Science Financière et d'Assurances	Directeur : M. Nicolas LEBOISNE
Institut National du Professorat et de l'Education	Administrateur Provisoire : M. Pierre CHAREYRON
Institut Universitaire de Technologie de Lyon 1	Directeur : M. Christophe VITON
Observatoire de Lyon	Directrice : Mme Isabelle DANIEL
Polytechnique Lyon	Directeur : M. Emmanuel PERRIN
UFR Biosciences	Administratrice provisoire : Mme Kathrin GIESELER
UFR des Sciences et Techniques des Activités Physiques et Sportives	Directeur : M. Yannick VANPOULLE
UFR Faculté des Sciences	Directeur : M. Bruno ANDRIOLETTI

REMERCIEMENTS

Tout d'abord, je remercie les membres du jury d'avoir accepté de lire et d'évaluer mon travail de doctorat. Merci de prendre le temps d'étudier un sujet pluridisciplinaire : j'espère être à la hauteur dans votre domaine d'expertise tout en étant pédagogue dans l'autre. Merci à **Peggy Cellier** et **Denis Pierron** de vous être rendu·e·s disponibles pour être rapporteur·e·s de cette thèse. Merci également à **Céline Rouveirol**, **Caroline Huart** et **Mohand-Saïd Hacid** d'avoir accepté d'endosser le rôle d'examinateur·rice.

Commençons par le commencement : je remercie **Fabien** qui m'a aiguillée vers Marc pour un stage M2. Sans lui, je n'aurais pas rencontré Marc, Moustafa et Céline, l'équipe DM2L ni même l'équipe Neuropop. Merci à mes **deux encadrants de thèse** et également à **Céline**, qui ont encadré mon stage de M2 et se sont battus pour trouver les financements nécessaires pour me lancer dans l'aventure du doctorat. Merci pour vos encouragements et motivations. Merci de croire en moi plus que moi-même. Merci **Moustafa** de m'avoir fait explorer un nouveau monde : les Neurosciences. Tu as été mon guide tout le long de cette aventure et tu as su m'expliquer toujours avec simplicité et aisance les différentes parties de ce monde. Merci de me supporter dans la médiation scientifique et je suis ravie d'avoir également collaboré avec toi sur des projets annexes à la thèse qui m'ont beaucoup apporté. Merci **Marc** d'avoir toujours répondu présent. Tu es toujours là pour arriver à la rescousse lorsqu'un doute plane ou qu'une interrogation se pose. Je te remercie également de m'avoir laissé la liberté de faire ce que je voulais dans tes cours. J'ai pu tester des idées farfelues, remettre en question ma vision de l'enseignement et je l'espère, améliorer ma pédagogie. Merci **Arnaud** d'avoir collaboré avec moi ces dernières années. Débarqué dans le domaine des Neurosciences en thèse n'a pas été facile pour moi, tu m'as énormément appris et j'espère que c'était un chouïa réciproque. Dans la série des superviseurs, je tiens à remercier également **Yanis**, mon chef lors de mon précédent emploi chez Aryballe qui a eu un impact très positif sur la perception que j'avais de moi-même sur le plan professionnel et personnel. Merci, je pense avoir énormément muri grâce à toi.

Malgré tout ce qu'ils m'ont apporté, les superviseurs ne suffisent pas pour réaliser une thèse. Je remercie donc également **toute l'équipe NEUROPOP** et tout particulièrement la team Tonic. **Juliette**, merci, pour ton précieux soutien : on a, certes, partagé beaucoup de Gin To mais également tous nos doutes et nos craintes concernant notre avenir et cela m'a beaucoup aidé. **Inès**, merci, tu es une personne magnifique et bien sûr, je ne parle pas (seulement) de ton physique : tu es la personne la plus bienveillante et forte que je connaisse. Tu ne lâches rien malgré tous les obstacles qui peuvent se dresser sur ton chemin. J'admire cette persévérance dont tu as fait preuve à de multiples reprises et je ne suis pas sûre qu'on serait allé au bout du concours « Créathon » sans toi. **Matthias**, merci mon roux préféré, grâce à toi, j'ai eu droit à une intégration en quelques jours plus que réussie. Confinée au bout d'un mois de thèse seulement, et pourtant, je ne me suis jamais sentie seule et tu n'y es pas pour rien. Merci **Morgane**, malgré notre grande différence d'âge, je me suis tout de suite sentie en confiance avec toi. Tu sais toujours dire les mots justes aux bons moments, tu es d'un soutien inestimable et j'admire ta maturité.

Je remercie également toutes les personnes qui ont partagé mon bureau et qui m'ont entendu insulter mon code plus d'une fois. Merci **Stéphane** et **Laura** pour ces débats toujours « fondamentaux » et « constructifs ». **Anna**, je suis contente qu'il n'y ait jamais eu de porte manteau dans notre bureau. Trêve de plaisanterie, j'aime beaucoup ta personnalité à la fois sensible et cash. Merci **Valentine**, ma coach en prononciation anglaise, je suis désolée de ne pas avoir pu poursuivre mes cours de python, on s'y remet bientôt ! **Tasha**, merci pour toutes tes petites attentions. **Alice**, tu es un adorable rayon de soleil ! J'espère que tu vivras bientôt cette formidable aventure qu'est la thèse.

Merci également, aux membres du bureau B04, **Erwan**, **Théo** et **Jules**, c'est toujours agréable de partager un repas ou une soirée avec vous. Vous êtes des travailleurs acharnés, bon courage pour la suite de la thèse ! Merci **Marylou**, j'ai beaucoup apprécié nos collaborations que ce soit sur des jeux de données philosophiques ou dans l'association de médiation scientifique Démesures. J'ai découvert plein de choses à tes côtés et j'espère que c'est réciproque. Merci **Catherine**, d'avoir pris le temps de relire plusieurs de mes travaux, merci pour tous tes retours toujours pertinents et pour ton humour quotidien. Merci **Camille** d'avoir pris le temps de partager avec moi ton parcours qui t'a permis de devenir l'admirable chercheuse que tu es aujourd'hui. Merci à **Christelle**, notre coach sportif, sécurité et expérimentation au labo, merci pour ton humour et ta bienveillance journalière. Merci aux post-doctorants **Halina** et **Christophe** d'être toujours présents pour aider les autres et pour partager vos fascinantes expériences. Merci **Aseel** et **Lucile** pour votre joyeuse humeur quotidienne, vous êtes chaleureuses et inspirantes. Merci à d'autres personnes du laboratoire comme **Hervé**, **Marc**, **Florence**, **Alexandra**, et **Romain** pour leur aide matérielle, technique ou administrative. Je remercie également toute l'équipe **DM2L** et quelques anciens, dont **Anes**, **Adnene**, **Romain** et **Tarek**.

Si j'en suis arrivé là, ce n'est pas seulement grâce à mes collègues, mais également à ma famille et mes amis. Je remercie donc **mon chéri**, **ma mère**, **mon père**, **ma belle-mère**, **mon frère** ainsi que **ma belle-famille**. Je remercie tout particulièrement **mon compagnon** qui ne comprend pas grand-chose à mon travail, mais qui fait son maximum pour m'aider lorsque je lui expose les difficultés scientifiques et méthodologiques que je rencontre quotidiennement. Je le remercie également pour son soutien lorsque je lui partage mes peurs et mes doutes. Je remercie également pour son soutien lorsque je lui partage mes peurs et mes doutes. Je remercie particulièrement **mes parents** qui m'ont aidé à m'orienter alors que je n'avais aucune idée de la suite de mon parcours. Merci d'avoir tous cru en moi.

Merci à mes amis. Je remercie **Hélène**, **Sabine**, **Élise**, **Marion**, la team **Tabouret**, et les **Ary'beaux** pour leur soutien ces trois dernières années. Je remercie particulièrement **Laura** qui a participé à la création de plusieurs figures de ce manuscrit, **Lohan** qui a géré le tournage de ma vidéo pour postuler à MT180 et **Solenne** qui a dessiné ma diapositive pour la finale régionale du concours.

Merci à tous les **membres de ComSciCon France**, j'adore organiser ce workshop de communication scientifique à vos côtés. Je tiens à remercier tout particulièrement **Mathilde**, **Antoine**, **Elodie** et **Marine**, vous êtes bienveillant·e·s et adorables et c'est toujours agréable de vous retrouver même lorsque nos réunions se terminent beaucoup trop tard.

Merci **Oriane**, vous avez été une coach professionnelle hors pair. Pour reprendre votre analogie du trajet en covoiturage : vous avez été une excellente copilote. Je sais maintenant quelles sont les destinations qui m'intéressent et je sais utiliser les outils à ma disposition pour m'orienter correctement.

Merci à tous mes **élèves** qui n'osaient pas lever la main au premier cours, mais qui la levaient tous aux derniers cours.

Merci également à tous les **participants** de mes expériences qui m'ont laissé explorer leur cerveau et leurs émotions. Une pensée toute particulière aux plus douillets.

Pour finir, je vais avoir l'audace de me remercier moi-même, car on ne va pas se mentir même si on communique avec pas mal de gens pendant une thèse, on travaille essentiellement seul. J'ai passé trois ans à essayer de faire des choses dont je n'avais aucune idée de comment elles se faisaient. La fierté n'est pas (encore) là, mais je me suis vu évoluer, je ne me suis jamais autant surpassée, j'ai tant appris et merci à tous de m'avoir soutenue et guidés, car le chemin aurait été impossible à parcourir seule. Merci à tous ceux qui ont croisé ma route ces trois dernières années.

AVANT-PROPOS

Cette thèse est à la frontière entre différentes disciplines : celles des Neurosciences et celle de la Data science. Soyez averti·e que ce manuscrit ne ressemble peut-être pas aux formats de présentation d'un travail de thèse dont vous avez l'habitude. Le format choisi ici est une thèse sur articles. Ce format se caractérise par une grande introduction, suivie d'un enchaînement d'articles qui constitue les différents chapitres de la thèse et se finit avec une discussion générale.

L'introduction générale présente les thématiques clefs abordées dans la thèse. Elle est suivie d'une partie « objectifs » décrivant les problématiques abordées et introduisant le contexte dans lequel s'inscrivent les différents articles. Les articles présentent les études menées au cours du doctorat pour atteindre les objectifs fixés. Ces articles peuvent être en cours d'écriture, soumis ou publiés. Ensuite, la discussion fait le lien entre les articles précédemment présentés : elle explique comment les différentes études réalisées répondent aux projets de recherche, compare leurs résultats, regroupe les différentes hypothèses formulées et propose des perspectives au travail effectué. Les articles s'adressent à des journaux internationaux, ils sont donc écrits en anglais, cependant tout le reste du manuscrit est écrit en français.

Il y a également une autre information que je souhaiterais vous partager afin que vous ne soyez pas surpris·e : la médiation scientifique faisant pour moi partie des rôles d'un chercheur, j'ai tenu à la mettre en avant, dans le manuscrit, en ajoutant certaines contributions de communication scientifique sous forme d'encadrés dans l'introduction. Ces encadrés ne sont pas nécessaires à la compréhension du manuscrit, vous pouvez les consulter ou non à votre guise.

Comme vous pouvez l'observer, j'essaie d'utiliser au mieux un langage inclusif. Nous n'y sommes pas forcément habitués, j'espère sincèrement que cette particularité ne vous troublera pas.

Je vous souhaite une agréable lecture et j'espère que vous refermerez ce manuscrit avec le sentiment d'avoir découvert des choses intéressantes !

TABLE DES MATIERES

REMERCIEMENTS.....	5
AVANT-PROPOS.....	9
RÉSUMÉ	13
ABSTRACT.....	15
LISTES DES ABRÉVIATIONS.....	17
INTRODUCTION	18
I. Les Émotions comme phénomène à la fois individuel et universel	19
A. Caractérisation de l'émotion	19
1. Définitions et théories des émotions.....	19
a) Théories	20
b) Définitions	23
2. Altération de l'émotion et de son expression	24
a) Déficit de l'expression des émotions	24
b) Troubles émotionnels.....	25
c) Déficit émotionnel	25
3. Fonctions	26
a) Chez l'adulte	26
b) Chez l'enfant.....	26
4. Universalité de l'émotion	27
B. Neurophysiologie de l'émotion	28
1. Mécanismes périphériques de l'émotion	28
2. Mécanismes centraux de l'émotion	29
a) Régions cérébrales de l'émotion.....	29
b) Réactions cognitives et expérientielles	30
c) Latéralisation corticale.....	30
C. Étude de l'émotion	31
1. Mesure du ressenti émotionnel.....	31
2. Mesure du comportement et des expressions faciales et vocales	32
3. Mesure de l'activité du système nerveux autonome.....	32
4. Mesures de l'activité du système nerveux central.....	33
a) IRM.....	33
b) EEG.....	34
II. Les odeurs comme vecteur d'émotions	36
A. L'olfaction : un sens sous-estimé	36
1. Pertes olfactives et qualité de vie : témoignages	36
2. Principales fonctions de l'odorat.....	37
B. De la molécule chimique à la perception de l'odeur	37
1. Odeur et odorant	37
2. Mécanismes périphériques de l'olfaction	39
a) Le mucus.....	40
b) L'épithélium olfactif.....	41
3. Mécanismes centraux de l'olfaction	42
a) Le bulbe olfactif.....	42
b) Le cortex olfactif primaire et secondaire	42
C. Le pouvoir émotionnel des odeurs	43
1. Langage	43
a) Langage courant.....	43

b) Description d'odeur.....	43
2. Odeurs, émotions et système nerveux.....	45
a) Système nerveux central	45
b) Système nerveux périphérique	46
D. La variabilité individuelle dans la perception olfactive.....	47
1. Génétique	47
2. Hormones et sexe biologique.....	47
3. Apprentissage.....	48
4. Environnement olfactif et culture	49
5. Conditions environnementales physique et chimiques	49
6. Tabagisme	50
7. Vieillesse	50
8. Maladies.....	50
9. Émotions	51
10. Familiarité	51
11. Appétit.....	51
III. L'intelligence artificielle comme outil d'analyse de la relation émotions et odeurs	53
A. Vocabulaire et définitions.....	53
B. Processus de découverte de connaissances dans les bases de données.....	54
C. Type de données de base	56
1. Données multidimensionnelles	57
2. Graphes	59
3. Itemsets	60
D. Méthodes utilisées	61
1. Découverte de sous-groupes	61
2. EMM.....	62
IV. Problématique et objectifs de la thèse	64
RÉSULTATS	66
ÉTUDE 1 : « Utilisation de la découverte de sous-groupes pour relier le caractère hédonique et l'intensité des odeurs aux réactions du système nerveux périphérique» (publiée)	67
ÉTUDE 2 : « Utilisation de l'extraction de sous-graphes attribués exceptionnels pour explorer la variabilité interindividuelle dans le traitement du caractère hédonique des odeurs au sein du cortex piriforme et l'amygdale» (en préparation).....	84
ÉTUDE 3 : « Explorer les caractéristiques temporelles du traitement neuronal des dimensions d'éveil émotionnel et de valence hédonique dans la modalité olfactive : une approche par data mining » (en préparation).....	108
ÉTUDE 4 : « Tutoriel d'extraction de connaissances à partir de données psychophysiologiques » (accepté).....	126
ÉTUDE 5 : « La science des données pourrait aider à tenir compte de la complexité et de l'hétérogénéité des données en sciences sensorielles et affectives » (publié).....	140
DISCUSSION	144
I. La science des données : un domaine d'analyse adéquat pour l'exploration du lien entre émotions et olfaction	145
A. La découverte de sous-groupes : un outil permettant de faire le lien entre l'olfaction et les émotions	145
B. La quantification de la participation des individus aux motifs : des mesures originales permettant de mieux comprendre la variabilité inter-individuelle	148
II. Représentation de la composante émotionnelle des odeurs au sein du système nerveux	149

A.	Émotions, odeurs et système nerveux	149
B.	Variabilité interindividuelle, odeurs et système nerveux	151
III.	Limites	153
A.	Limites expérimentales.....	153
1.	Nombre d'individus et d'odeurs.....	153
2.	Facteurs de variabilité.....	154
3.	Données utilisées	154
B.	Limites méthodologiques	155
IV.	Perspectives.....	155
A.	Approches développées prometteuses et polyvalentes.....	155
B.	La compréhension du lien odeurs et émotions : créatrice de nouvelles opportunités	156
V.	Le mot de la fin	157
RÉFÉRENCES		158
ANNEXE.....		174
	Annexe 1 : « Étude de la variabilité de perception des odeurs à partir de descriptions subjectives » (soumis).....	175
	Annexe 2 : « Perte de l'odorat avec la COVID-19 et impact sur la qualité de vie » (publié)	200
	Annexe 3 : « Perte de l'odorat comme indicateur précoce de l'évolution de la pandémie de COVID-19 » (publié).....	213
	Annexe 4 : « Résultats supplémentaires de l'Étude 1 » (publié).....	222
	Annexe 5 : « Résultats supplémentaires de l'Étude 2 » (en préparation)	241

RÉSUMÉ

La dimension subjective et émotionnelle des odeurs est une composante majeure de la perception olfactive. Elle déclenche des comportements d'approche ou d'évitement essentiels à la survie de l'individu en participant à la prévention de dangers potentiels, à notre alimentation et à la communication sociale. Aujourd'hui cependant, la façon dont ces affects olfactifs se manifestent au niveau du système nerveux n'est encore pas entièrement comprise. Quant à la grande variabilité inter-individuelle dans la perception olfactive, si elle n'est plus à démontrer, elle est à l'heure actuelle peu explorée en neurophysiologie de l'émotion olfactive, faute d'outils d'analyse adéquats.

Dans ce projet de thèse, nous avons cherché à lever ces verrous méthodologiques par des approches provenant de l'intelligence artificielle. Dans ce cadre, des méthodes de science des données ont été utilisées dans le but de comprendre comment la diversité émotionnelle en matière d'odeurs se reflète au niveau neurophysiologique.

Nous nous sommes spécifiquement intéressés à l'exploration de modèles exceptionnels et la découverte de sous-groupes. Ces méthodes descriptives permettent d'extraire des informations précises sur le comportement commun à un groupe d'individus, mais également de nous donner des indications à l'échelle individuelle. Ainsi, nous les avons appliquées à différentes mesures neurophysiologiques de l'émotion : des mesures du système nerveux autonome et des mesures du système nerveux central (IRMf et EEG) à travers trois études expérimentales.

Dans la première étude, nous avons décrit le caractère hédonique (plaisant/déplaisant) et l'intensité de l'odeur par des réactions physiologiques spécifiques. Dans la seconde étude, nous avons localisé spatialement les composantes agréable et désagréable de l'odeur dans le cortex piriforme et l'amygdale. Dans la troisième étude, nous avons exploré la dimension temporelle de ce traitement hédonique.

L'utilisation de ces approches de fouille de données se révèle particulièrement adaptée, car les modèles descriptifs qui en ressortent sont en accord avec les études passées. En effet, l'étude 1 soutient l'hypothèse d'une réponse somatique plus forte pour les stimuli olfactifs aversifs (désagréables et intenses). L'étude 2 corrobore avec les études démontrant un traitement asymétrique des odeurs désagréables au niveau du cortex piriforme antérieur. L'étude 3 montre une réponse tardive pour les odeurs plaisantes et un traitement précoce pour les odeurs déplaisantes, déjà constatés auparavant. Au-delà de ces résultats, qui confirment la littérature antérieure, nous avons mis en évidence des règles liant comportement hédonique et activité nerveuse, notamment asymétrique au niveau hémisphérique, ou en lien avec le caractère invariant de la réponse hédonique entre individus. Ces résultats, qui n'avaient pas été révélés précédemment, ouvrent la porte à de nouvelles études liant odeur et émotions. Ces approches nous permettent en effet d'explorer la diversité individuelle présente pour chaque description trouvée. L'étude de la variabilité indique une plus grande hétérogénéité pour les odeurs agréables que pour les odeurs désagréables concernant l'activité du système nerveux périphérique (étude 1) et la localisation cérébrale (étude 2).

Cette tendance s'inverse lorsque l'on considère la dimension temporelle. En effet, les odeurs agréables entraînent une même réponse bien précise (activité bilatérale tardive dans les basses fréquences) alors que les individus donnent des réponses différentes.

Afin de permettre aux non-spécialistes du domaine de la data science d'appliquer ce type d'approche à leurs propres données, nous avons proposé, dans une quatrième étude, un protocole méthodologique. Ce tutoriel décrit étape par étape l'application de la découverte de sous-groupes sur des données de conductance cutanée.

Nous avons également réalisé un travail théorique dans une cinquième étude. Il offre une réflexion sur l'analyse des émotions en condition réelle hors des laboratoires. Il propose ainsi des méthodes provenant de la science des données comme solution possible pour analyser les données nombreuses et hétérogènes qui en résultent.

Prises ensemble, ces données renforcent la découverte de réseaux neuronaux différents impliqués dans les comportements d'approche et de retrait en matière d'odeurs. Elles suggèrent également qu'une approche d'exploration de données peut être d'une grande utilité pour développer des modèles descriptifs et explicatifs de l'émotion et de l'olfaction.

Mots clefs : olfaction, émotion, neurophysiologie, data mining, découverte de sous-groupes

ABSTRACT

The subjective and emotional dimension of odors is an important component of olfactory perception. It enables approach or avoidance behaviors essential to the survival of the individual. Indeed, it participates in the prevention of potential dangers, in nutrition and in social communication. The way in which these olfactory effects are manifested in the nervous system is still not understood fully today. Moreover, although the high inter-individual variability in olfactory perception no longer needs to be demonstrated it remains little explored at the present time. In fact, due to the lack of adequate analysis tools, this variability is not taken into account in studies concerning the neurophysiology of emotion in olfactory perception.

In this thesis project, we sought to remove these methodological barriers using approaches from artificial intelligence. In this framework, data science methods were used with the aim of understanding how emotional diversity in odors is reflected at the neurophysiological level.

We were specifically interested in exploring exceptional patterns and subgroups discovery. These descriptive methods allow us to extract precise information about the common behavior of a group of individuals, but also to give us insights at the individual level. We applied them to different neurophysiological measures of emotion: autonomous nervous system measures and central nervous system measures (fMRI and EEG) through three experimental studies.

In the first study, the hedonic character (pleasant/unpleasant) and intensity of the odor were described by specific physiological reactions. In the second study, the pleasant and unpleasant components of the odor were localized spatially in the piriform cortex and amygdala. In the third study, the temporal dimension of this hedonic treatment was explored.

The use of these data mining approaches is particularly suitable because the descriptive patterns that emerge are consistent with past studies. Indeed, study 1 supports the hypothesis of a stronger somatic response to aversive (unpleasant and intense) olfactory stimuli. Study 2 is in agreement with studies showing asymmetric processing of unpleasant odors in the anterior piriform cortex. Study 3 shows a delayed response for pleasant odors and an early treatment for unpleasant ones which has already been demonstrated. Further to these results, which confirm the previous literature, we have proposed rules linking hedonic behavior and nervous activity, notably asymmetric, or related to the invariant character of the hedonic response between individuals. These results, which had not been revealed before, open the door to new studies linking odor and emotions. Indeed, these approaches allow us to explore the individual diversity present for each description found. The study of variability indicates a greater heterogeneity for pleasant odors than for unpleasant odors concerning the activity of the peripheral nervous system (study 1) and the cerebral localization (study 2). This trend is reversed when considering the temporal dimension. Indeed, pleasant odors create a very precise response (late bilateral activity in the low frequencies) whereas different responses are observed according to the individuals exposed to unpleasant odors.

In order to allow non-specialists in the field of data science to apply this type of approach to their own data, we have suggested, in a fourth study, a methodological protocol.

This tutorial describes step-by-step the application of subgroup discovery on skin conductance data.

We also carried out a theoretical work in a fifth study. This study looks at the analysis of emotions in real conditions outside the laboratory. It suggests data science methods provide a possible solution to analyze the voluminous and heterogeneous data produced.

Taken together, these data reinforce the idea of different emotional networks involved in odor approach and withdrawal behavior. They also suggest that using a data mining approach can be of great help in developing descriptive and explanatory models of emotion and olfaction.

Keywords: olfaction, emotion, neurophysiology, data mining, subgroup discovery

LISTES DES ABRÉVIATIONS

AMPc	Adénosine monophosphate cyclique
Amyg	Amygdale
c.-à-d	c'est-à-dire
PC	Cortex piriforme
EEG	Électroencéphalographie
EMM	Exploration de modèles exceptionnels
Entho C	Cortex entorhinal
Hipp	Hippocampe
IA	Intelligence artificielle
ID	Identité
IRM	Imagerie par résonance magnétique
IRMf	IRM fonctionnelle cérébrale
NSSCR	Conductance cutanée non spécifique
OFC	Cortex orbitofrontal
PA	Potentiel d'action
PET	Tomographie par émission de positons
ex.	Par exemple
SC	Conductance de la peau
SNC	Système nerveux central
SNP	Système nerveux périphérique
Thal	Thalamus

INTRODUCTION

L'ambition de cette thèse est double. Tout d'abord, nous souhaitons apporter des éléments de connaissance sur les relations entre odeurs et émotions. L'odeur est un vecteur d'émotions. Ces émotions peuvent être partagées entre les individus mais en fonction des trajectoires développementales de chacun, ces états affectifs peuvent également être spécifiques à une personne ou à un groupe de personnes. Notre première ambition est de comprendre comment le cerveau humain représente ces états affectifs induits par les odeurs et aussi comment la variabilité interindividuelle s'exprime à ce même niveau neuronal. Pour répondre à ces questions, il faut des outils et des approches capables de mesurer les réponses neuronales aux odeurs chez l'homme mais aussi des moyens d'analyse qui prennent en compte la variabilité interindividuelle décrite ci-dessus. La seconde ambition de cette thèse est d'intégrer au processus classique d'analyse des données neuroscientifiques, une analyse des données issue de l'intelligence artificielle et en particulier des sciences des données. Cette introduction est structurée en quatre parties. Premièrement, nous présenterons différents modèles de l'émotion. Dans un deuxième temps, nous détaillerons la modalité olfactive dans toute son universalité et sa diversité. Dans une troisième partie, nous définirons les approches utilisées en intelligence artificielle et en particulier en data sciences. Enfin, dans une quatrième partie, nous décrirons les différents objectifs spécifiques de la présente thèse de doctorat ainsi que les différentes études mises en place pour tester ces objectifs.

I. Les Émotions comme phénomène à la fois individuel et universel

Dès l'antiquité, les philosophes grecs et romains ont analysé les émotions [1,2]. Depuis plus de 2000 ans, de nombreux penseurs et scientifiques, tels que Platon et Aristote [3,4], Descartes [5,6], Darwin [7], Freud [8,9], William James [10,11] et Walter Cannon [12], ont débattu sur la nature et le fonctionnement des émotions. Ce sujet fait encore l'objet de controverses scientifiques. Nous allons donc voir les différents points de vue existants à ce sujet et celui que nous comptons adopter dans le cadre de cette thèse. Nous verrons également les fonctions et altérations de l'émotion. Ensuite, nous pourrons entrevoir sa dimension universelle. Pour finir, nous détaillerons la neurophysiologie de l'émotion et nous aborderons les différentes manières de la mesurer.

A. Caractérisation de l'émotion

1. Définitions et théories des émotions

Le terme émotion est souvent utilisé de manière interchangeable avec d'autres termes comme la sensation, le sentiment, l'humeur ; ces termes ne sont cependant pas synonymes.

La sensation est une perception immédiate déclenchée par nos récepteurs sensoriels (via le toucher, la thermoception ou la douleur, par exemple) [13,14].

Le sentiment renvoie à la composante cognitive et expérientielle de l'émotion. Un sentiment n'a pas de dimension physiologique (*voir encadré 1*), comportementale et expressive [13].

Le tempérament est un trait de caractère inné qui joue un rôle, entre autres, sur le seuil de déclenchement des émotions, sur leur intensité et leur régulation [15–17].

Contrairement à l'émotion, l'humeur est un état qui apparaît lentement et son déclenchement n'est pas facilement identifiable. Elle se trouve en arrière-plan par rapport à nos pensées et se manifeste moins intensément, plus souvent et plus longtemps (quelques heures ou jours) par rapport à l'émotion qui est rare, brève (quelques secondes ou minutes), intense et focalisée sur un objet, un déclencheur précis. L'humeur et les émotions sont tout de même liées : l'humeur influe sur le seuil de déclenchement des émotions, leur intensité, leur durée et leur contrôle. Un individu va, par exemple, plus facilement se mettre en colère s'il est d'humeur irritée que s'il est d'humeur joyeuse. Une émotion très intense et récurrente peut laisser place à une humeur. De même, une émotion non exprimée peut laisser place à une humeur négative [13,18,19].

Encadré 1. Système nerveux

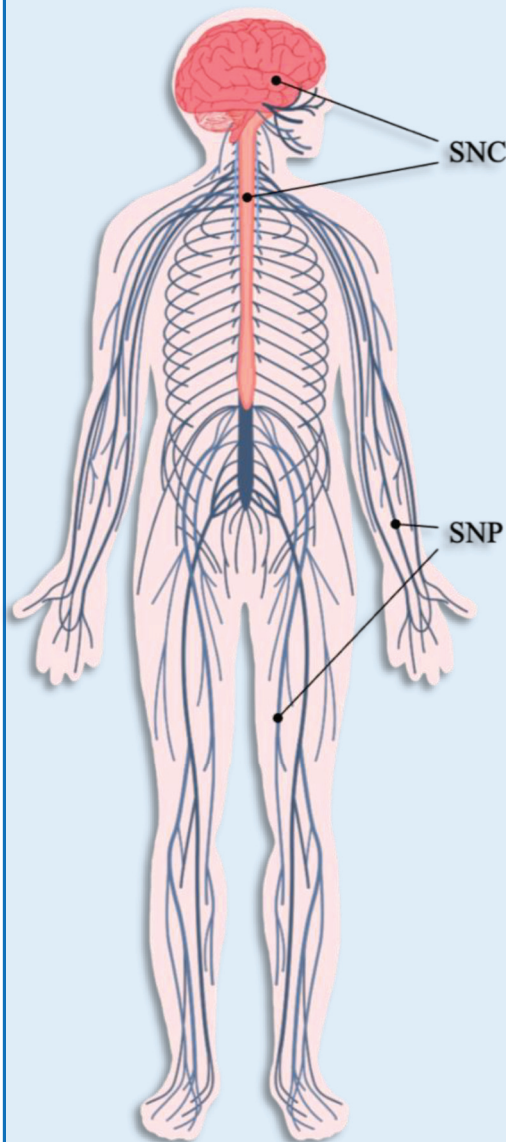


Figure 1 - Système nerveux chez l'humain

Notre système nerveux est regroupé en deux structures anatomiques différentes appelées système nerveux central et système nerveux périphérique.

L'encéphale, et la moelle épinière forment le système nerveux central (SNC). L'encéphale est composé du cerveau, du cervelet et du tronc cérébral. La moelle épinière (ou moelle spinale) est présente le long de la colonne vertébrale, elle transmet les messages nerveux entre l'encéphale et le reste du corps et est également responsable de certains réflexes.

Les nerfs qui partent de la moelle épinière appelé nerfs spinaux (ou nerfs rachidiens), ainsi que les nerfs qui partent du tronc cérébral nommés nerfs crâniens et les ganglions situés hors du SNC composent le système nerveux périphérique (SNP). Ces nerfs font le lien entre le SNC et les récepteurs sensoriels mais également entre le SNC et les muscles.

Le SNP intervient dans le système nerveux somatique qui permet la perception des stimulations externes (système sensitif) et le contrôle des mouvements volontaires ou réflexes via les muscles squelettiques (système moteur).

Le SNP intervient également dans le système nerveux autonome: il coordonne de façon non volontaire le fonctionnement des organes internes dont le système digestif, et cardiovasculaire. Il peut, par exemple, préparer un individu à une action de fuite en augmentant son rythme cardiaque et en stoppant la digestion. Cette manifestation corporelle est nommée réaction physiologique [20].

a) Théories

Nous savons tous ce qu'est une émotion et pourtant il est très difficile d'en donner une définition. De nombreux scientifiques ont étudié l'émotion et ont tenté de la décrire. Certain·e·s s'intéressent aux mécanismes de déclenchement des émotions, d'autres à la chronologie des différents processus présents dans l'émotion et d'autres encore aux mécanismes de différenciation entre les différentes émotions. Cela a donné lieu à de nombreuses théories dont les plus connues et contemporaines sont décrites ci-après.

William James [11] et Carl Lange [21] proposent indépendamment une théorie similaire appelée la **théorie périphérique des émotions** [10]: chaque émotion posséderait

son propre pattern d'activation du système nerveux périphérique (*voir encadré 1*). C'est la perception des changements physiologiques résultant de cette activité qui déclencherait l'émotion. Ainsi nous aurions peur car – par exemple – notre rythme cardiaque s'accélère, nous transpirons et nous tremblons.

Cannon et Bard proposent une théorie assez différente appelée **théorie centrale des émotions** [12,22]. Pour eux, le déclenchement d'une émotion est déterminé non pas au niveau périphérique mais au niveau du système nerveux central, et plus particulièrement par le traitement du stimulus au sein du thalamus. L'excitation du thalamus induirait simultanément un changement physiologique et un ressenti émotionnel.

Schachter et Singer ajoutent l'évaluation de la situation aux théories précédentes, et proposent la **théorie bi-factorielle** [23,24]. Une émotion est déterminée par la présence simultanée de deux facteurs : une réaction physiologique et une activité cognitive en lien avec la situation déclenchante de cette activation physiologique. L'interprétation de la situation dépend des expériences passées de l'individu et détermine l'émotion ressentie. Ainsi lorsqu'un individu perçoit un stimulus émotionnel, il établit un lien entre son activité physiologique et une explication pertinente à cette activité en fonction de la situation présente. C'est ce mécanisme qui déclencherait l'émotion.

En outre, plusieurs scientifiques (Lazarus [25], Sander, Grandjean & Scherer [26], Siemer, Mauss & Gross [27]) suggèrent que l'évaluation cognitive du stimulus est à l'origine du déclenchement de l'émotion. L'évaluation cognitive analyse de façon automatique et non consciente un certain nombre de critères comme la nouveauté de la situation, le caractère agréable/désagréable, les conséquences de l'induction de l'état émotionnel etc. Le résultat de la combinaison de ces évaluations est à l'origine d'une émotion spécifique. L'interprétation de l'évènement étant subjective, il existe une grande variabilité de réactions émotionnelles face au même évènement. Ainsi, d'après les **théories de l'évaluation cognitive**, si un individu évalue une menace, il va ressentir l'émotion de peur et les réactions physiologiques résultantes.

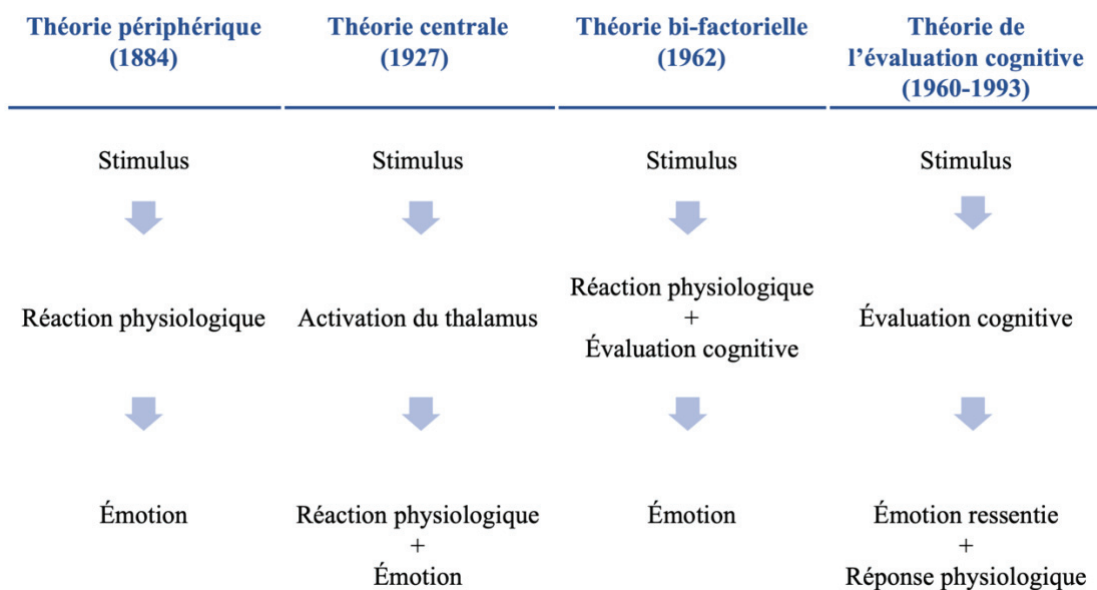


Figure 2 - Résumé des théories concernant les mécanismes de déclenchement des émotions [28,29]

Matsumoto et Ekman proposent la théorie des **émotions de base** [19,30,31]. Les émotions de base ou émotions primaires seraient universelles, en nombre limité et auraient chacune un rôle adaptatif dans l'évolution. Ces émotions comprennent la joie, la peur, la colère, la tristesse, le dégoût et la surprise. Chaque émotion de base a des caractéristiques qui lui sont propres : des conditions spécifiques d'émergence (ex. enterrement, mariage), un pattern d'activité physiologique spécifique, un pattern d'expression spécifique (ex. sourire) ainsi qu'une activation cérébrale spécifique. Par analogie avec les couleurs, les émotions primaires existent avec différentes intensités et peuvent se mélanger afin de donner un spectre complet de l'expérience émotionnelle humaine. Ainsi, une faible colère serait de la contrariété et une très forte colère conduirait ce qu'on appelle communément un état de rage. Une émotion secondaire comme la honte serait un mélange entre du dégoût de soi-même et de la peur. Plutchik a tenté d'illustrer cette théorie avec 8 émotions de base sous la forme d'une roue des émotions [32] illustrée Figure 3.

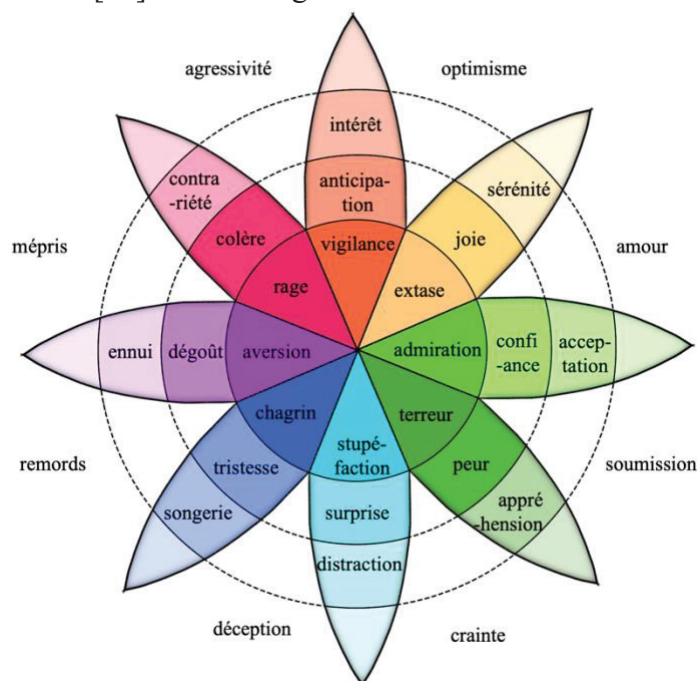


Figure 3 - Roue des émotions selon Plutchik. Chaque couleur correspond à une émotion de base. Les différentes intensités d'une couleur donnent les différents degrés de l'émotion (plus une couleur est foncée, plus l'émotion associée est forte). Entre chaque émotion se trouve une émotion plus complexe correspondant à la combinaison des deux émotions adjacentes.

Russell et Barrett proposent une modélisation **bi-dimensionnelle** des émotions [33]. En effet, ces dernières seraient décrites par deux dimensions élémentaires : l'éveil (ou arousal) et la valence. L'éveil (émotionnel) correspond au caractère calmant ou stimulant et la valence est assimilée au plaisir ou déplaisir ressenti. Un modèle circulaire (Figure 4) nommé « circumplex » a été proposé pour représenter l'emplacement de chaque émotion de base par rapport à ces deux axes. Ainsi la tristesse serait une émotion désagréable et peu stimulante et la peur une émotion stimulante et désagréable.

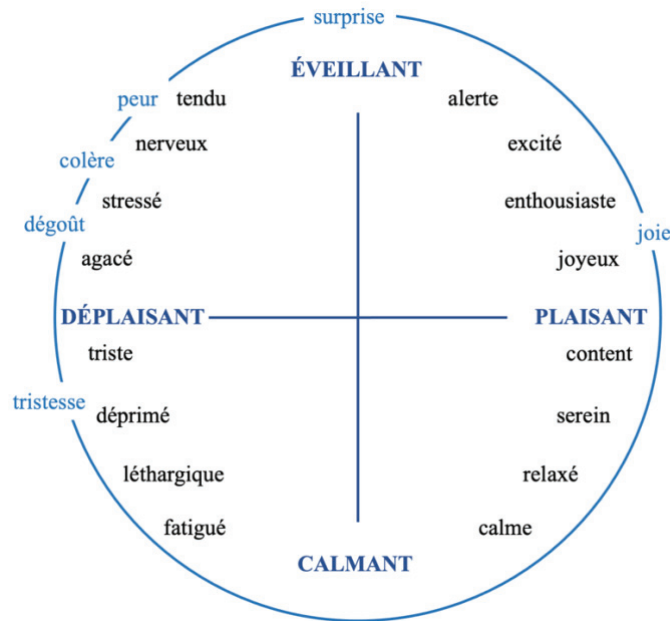


Figure 4 - Modèle circumplex de Russell. Chaque émotion serait décrite sur deux dimensions indépendantes : la valence (plaisir/déplaisir) et le degré d'éveil. Les émotions proches sur la structure circulaire sont corrélées positivement et celles placées à 180° sont corrélées négativement. Les émotions de base sont notées en bleu clair et les autres en noir.

Chacune de ces théories n'est pas suffisante à elle seule [28,29]. Elles se complètent et présentent, chacune, différentes composantes de l'émotion.

b) Définitions

Dans le cadre de ce travail de thèse, nous abordons l'étude des émotions en considérant toutes les théories présentées précédemment.

Une émotion est déclenchée par un évènement soudain qui interrompt nos pensées et notre activité, et focalise notre attention. Une réaction émotionnelle est brève et intense et se caractérise par 3 types de réactions [34–36] :

- 1) **Des réactions cognitives et expérientielles.** En effet, tout d'abord la situation doit être jugée importante pour soi afin de déclencher un traitement émotionnel. Par exemple, si pour un élève « réussir dans sa vie professionnelle » n'a pas d'importance, alors, généralement, l'échec à un examen ne déclenchera pas d'émotion négative sur lui [37]. Ainsi, une évaluation cognitive a lieu et en fonction de la nouveauté de la situation, de sa valence, de la soudaineté de l'évènement, de son degré de contrôle, de la compatibilité avec ses normes, une émotion spécifique peut être induite [38,39].

Cet évènement va également s'accompagner d'une expérience subjective ; l'individu va tirer des leçons sur ce qui s'est passé. Par exemple, une personne peut décider de ne plus jamais regarder son téléphone en conduisant, si elle manque d'écraser quelqu'un sur la route en écrivant un message. Par conséquent, les évènements émotionnels représentent une source d'information précieuse permettant d'élaborer des plans futurs afin d'éviter la reproduction de l'évènement dans le cas d'émotions

déplaisantes ou au contraire afin de favoriser les conditions de sa réapparition dans le cas d'émotions plaisantes.

- 2) **Des réactions physiologiques.** L'évènement émotionnel a un effet immédiat sur le corps (sudation, augmentation du rythme cardiaque, respiration saccadée ...) de par l'activation du système nerveux autonome (*voir encadré 1*) [40].
- 3) **Des réactions comportementales et expressives.** L'émotion se traduit également par une expression faciale (sourire, haussement de sourcils, yeux écarquillés) ou vocale (voix qui se casse, monte, ou tressaute), un changement de comportement (agitation, immobilisation) et une tendance à l'action (approche, fuite, violence) [13].

Une émotion se caractérise donc par un pattern d'évaluations cognitives, d'expériences subjectives, de mimiques, de tendances à l'action, et de réactions physiologiques. Ainsi, ces différentes réactions dépendent de l'émotion ressentie et plus généralement de la valence et du niveau d'éveil ou d'excitation en lien avec cette émotion.

2. *Altération de l'émotion et de son expression*

Afin de mieux appréhender le rôle des émotions, nous allons entrevoir quelques cas d'altération de l'émotion et découvrir les conséquences engendrées dans la vie des individus qui en sont victimes.

a) *Déficit de l'expression des émotions*

Certains individus ont perdu la capacité d'exprimer leurs émotions en modulant leur voix : ils sont atteints d'**aprosodie**. Ils éprouvent des expériences émotionnelles normales et pourtant ils parlent d'un ton monotone indépendant de leurs ressentis. Par exemple, un enseignant souffrant d'aprosodie aura du mal à maintenir le calme en classe du fait qu'aucune émotion de colère ou de tristesse ne se dégagera de son discours [41,42].

Il existe également une impossibilité à mettre des mots sur ses émotions. Dans le cadre de ce trouble, appelé **alexithymie**, la personne n'est pas capable de partager ses sentiments avec autrui mais également d'identifier ses propres sentiments. Seules les réactions corporelles sont perçues et peuvent être décrites. Les conséquences de cette incapacité sont un vocabulaire affectif très pauvre et une imagination, dépourvue de fantasmes avec une difficulté à raconter ses souvenirs ou ses rêves. Les descriptions d'évènements sont factuelles et centrées sur les circonstances. L'alexithymie n'est pas forcément pathologique, elle peut être privilégiée par certaines cultures où exprimer ses émotions est mal perçu. Elle peut également être passagère et intervenir dans une situation précise notamment lors d'un évènement traumatisant [43].

b) Troubles émotionnels

Une déviation de la réponse émotionnelle contribue à une désadaptation psychologique. Certains troubles sont caractérisés par la surexpression d'une émotion spécifique : la dépression par des émotions négatives, les troubles anxieux par des peurs excessives, la phobie sociale par une gêne et une honte extrême, le trouble de la personnalité antisocial ou les troubles d'extériorisation par une colère excessive et une gêne réduite.

De plus, beaucoup de troubles psychiatriques, comme la schizophrénie ou l'anorexie, impliquent des troubles émotionnels appelés troubles affectifs [44,45]. Un symptôme présent dans la dépression, la schizophrénie et les addictions, par exemple, est l'**anhédonie**, c'est-à-dire l'incapacité à éprouver du plaisir.

Par ailleurs, les maladies neurologiques peuvent également être associées à des troubles émotionnels. Dans les maladies d'Alzheimer et de Parkinson, des symptômes présents également chez les schizophrènes et dépressifs, apparaissent, comme l'**émoussement affectif** qui correspond à une diminution ou absence de réactions affectives [46].

c) Déficit émotionnel

Étudions maintenant la situation d'un individu qui aurait perdu une grande partie de ses capacités émotionnelles. C'est le cas de *Phineas Gage* dont une barre de fer a traversé le crâne durant une explosion ce qui a engendré une lésion au niveau du lobe frontal [6,47]. Outre le fait étonnant qu'aucune autre séquelle physique ou mentale (attention, perception sensorielle, langage, mémoire et intelligence intactes) n'a été détectée, la personnalité qu'il avait avant l'accident avait considérablement changé [48,49].

Avant l'accident Phineas Gage était décrit comme fin, équilibré ; il savait faire des choix qui pouvaient améliorer sa condition, avait le sens des responsabilités, se souciait de la qualité de son travail, était habile en affaires, persévérant, adapté socialement et respectait les règles morales dans ses relations avec les autres.

Après l'accident, il ignorait la « morale », se montrait grossier, montrait peu de respect pour ses amis, ne se préoccupait plus de son avenir, ne montrait aucun signe de prévoyance, prenait des décisions sans modération et clairement désavantageuses, supportait difficilement les contraintes ou conseils qui entravaient ses désirs et se comportait comme un enfant capricieux et inconstant.

Ce déficit de la perception émotionnelle semblait avoir changé son caractère, ses goûts, ses rêves et ses ambitions. Cela a entraîné une perte d'indépendance et des échecs sociaux et professionnels successifs.

Le cas de Phineas Gage n'est pas un exemple isolé, d'autres cas similaires ont été répertoriés [50–53]. Ces personnes porteuses de lésions préfrontales, semblent également avoir perdu leurs facultés émotionnelles en même temps que la capacité de programmer leurs actions dans l'avenir, de se conduire en fonction des règles sociales qu'ils avaient apprises et de faire des choix susceptibles d'être le plus avantageux pour leur survie [54].

3. Fonctions

A la lecture du paragraphe précédent, on comprend que les émotions sont vitales aussi bien d'un point de vue social que psychologique. Ici, nous allons distinguer les fonctions de l'émotion chez l'adulte d'une part et chez l'enfant d'autre part.

a) Chez l'adulte

Les émotions sont indispensables pour notre survie comme l'atteste le rôle qu'elles jouent dans les domaines suivants [45,55] :

- 1) **Réaction adaptative.** La physiologie de l'émotion aide à préparer les individus à une action rapide en réponse à un événement donné [56]. La peur, par exemple, permet de réagir face à une menace pour son intégrité par la fuite ou le combat [57].
- 2) **Comportement rationnel.** L'expérience de l'émotion sert à organiser les processus cognitifs. Elle nous indique la meilleure direction (ou la moins pire), nous permet de réaliser des projets en prenant des décisions dans notre intérêt et en réalisant des tâches (parfois déplaisantes) qui consistent à prévoir un avenir incertain et à programmer nos actions en conséquence. Elle nous permet ainsi de faire des jugements moraux et de mettre en œuvre des principes logiques [6,58].
- 3) **Interaction sociale.** L'expression faciale et vocale de l'émotion ainsi que la posture communiquent des informations importantes sur notre état d'esprit à l'entourage qui va pouvoir réagir en conséquence [59–61]. Mal interpréter une mimique peut nous faire adopter un comportement inadapté à la situation.
- 4) **Apprentissage.** Les émotions agissent sur la capacité de mémorisation et l'attention [62,63].

b) Chez l'enfant

Les facultés émotionnelles mentionnées ci-dessus pour l'adulte ne sont pas toutes nécessairement présentes chez l'enfant dont le cerveau est en pleine maturation. Effectivement, les structures du cerveau liées à l'expression et la compréhension émotionnelle se développent tout le long de l'enfance et de l'adolescence [64].

L'émotion joue un rôle dans le développement de l'enfant [65]. La compréhension et l'expression de certaines émotions pour la première fois émergent en même temps que d'autres capacités. Par exemple, la peur apparaît en même temps que le développement moteur et l'embarras apparaît en même temps que la conscience de soi [55,66].

De plus, les manifestations émotionnelles des parents sont une source essentielle d'informations sur l'environnement pour l'enfant [55]. En effet, leurs expressions faciales et vocales le récompensent pour certains comportements et le découragent pour d'autres plus dangereux comme s'approcher d'un précipice.

4. Universalité de l'émotion

Nous avons vu qu'une émotion est un phénomène subjectif du fait de notre réaction cognitive et expérientielle individuelle. Ainsi un même stimulus peut déclencher des émotions très différentes d'un individu à l'autre. Cependant, il peut être intéressant de se demander si une même émotion s'exprime ou non de la même manière chez tout le monde.

Les émotions pourraient avoir une composante culturelle. Le lexique des émotions est effectivement assez différent d'une culture à l'autre [67]. Certaines langues sont dépourvues de mots désignant deux états émotionnels particuliers : par exemple, les Maya Yucatec n'ont pas de mot permettant de discerner la colère et le dégoût [68], les indiens Rarámuri ne distinguent pas la culpabilité et la honte [69]. À l'inverse des émotions décrites par un mot unique dans notre langue comme la joie, sont représentées par de multiples émotions dans d'autres langues : il existe par exemple « *Nam jai* » en thaïlandais qui décrit le plaisir d'un sacrifice désintéressé ou « *Duende* » en espagnol qui traduit une émotion passionnelle induite, par exemple, par la beauté d'une œuvre d'art [70]. Le fait d'utiliser un vocabulaire différent concernant la dénomination des émotions ne signifie pas qu'il n'y a pas d'émotions humaines universelles mais nous n'avons aucune garantie que les émotions de base telles que la colère et la peur soient des étiquettes universelles pour catégoriser les émotions et que celles-ci soient ainsi naturellement et biologiquement fixées [71,72].

D'ailleurs, chaque émotion se caractériserait par une réponse physiologique spécifique [73]. Une revue de la littérature sur l'activité du système nerveux autonome dans l'émotion compare les réponses physiologiques observées dans 134 études pour un grand nombre d'émotions [74]. Beaucoup d'études font consensus bien que certaines soient en désaccord. Dans le cas du plaisir, par exemple, certaines études trouvent une augmentation du rythme respiratoire et d'autres une diminution. Certaines discordances sont expliquées par une différence dans les types de stimuli utilisés : pour le dégoût, par exemple, on observe généralement une activité électrodermale accrue, cependant cette réponse sera courte pour les images exprimant du dégoût et de longue durée pour les odeurs dégoûtantes [74]. Néanmoins toutes les incohérences des modèles de réponses relevées par les diverses études ne sont pas expliquées, nous ne savons donc pas si cela est dû à la présence de variabilités individuelles ou aux conditions expérimentales divergentes.

De la même façon, chaque émotion pourrait être identifiée par un groupe d'expressions et d'actions caractéristiques et distinctes des autres émotions. Par exemple, la joie provoquerait une expression faciale qui lui est propre : les sourcils intérieurs sont levés, la paupière inférieure resserrée, les joues surélevées, la lèvre supérieure relevée et les coins des lèvres tournés vers le haut [75]. Les études interculturelles de Paul Ekman et Carroll Izard ont démontré que ces expressions faciales de l'émotion sont universelles [76,77]. Les individus de cultures différentes sont capables d'identifier avec précision les manifestations faciales de colère, de dégoût, de peur, de joie, de tristesse, de surprise, de honte, de mépris, d'embarras et d'amusement. De plus, ils ont, chacun, un comportement émotionnel similaire lorsqu'ils sont seuls mais peuvent montrer des manifestations émotionnelles très différentes en présence d'une figure d'autorité [78]. Ainsi nous aurions naturellement la même manière

d'exprimer nos émotions mais nous les réprimerions en fonction du contexte social et culturel.

Nous pouvons observer l'universalité de l'émotion à une échelle plus grande encore en considérant les autres espèces animales. En effet, on observe la présence d'émotions de base chez de nombreux animaux dont les primates ou les animaux domestiques [79]. Les émotions chez les animaux peuvent être appréhendées comme chez l'humain : elles se manifestent également par des réponses cognitives, physiologiques, comportementales et expressives [7,80–82]. Si la composante physiologique et comportementale chez un non humain est mesurable de la même façon que pour un humain, il n'est pas possible de demander un rapport verbal à l'animal pour accéder à ses expériences émotionnelles et leurs antécédents. La composante cognitive est donc évaluée d'une autre manière [83]. L'évaluation cognitive et les réactions expérientielles peuvent être déduites en fonction du changement comportemental et physiologique en réponse à des situations pour lesquelles les critères d'évaluation (tels que la soudaineté, la familiarité et la prévisibilité) ont été manipulés expérimentalement. Les relations directes entre les évaluations présumées et les résultats émotionnels mesurables peuvent ensuite être établies. De cette façon, les scientifiques peuvent observer qu'il y a un déclenchement de réactions adaptées au stimulus et que de nouvelles réactions se sont installées en réponse aux stimuli précédemment rencontrés chez certaines espèces comme la truie et l'agneau [84,85].

L'émotion est importante pour la communication au sein de l'espèce [86] mais également pour la communication avec d'autres espèces. À notre échelle d'humain, au fur et à mesure de la domestication, nous avons appris à interpréter l'état émotionnel des animaux à partir de leur comportement afin d'adapter nos actions en conséquence. On évite, par exemple, de caresser un chien ou un chat qui semble énervé. De façon réciproque, les animaux perçoivent également nos émotions de par nos comportements, nos expressions vocales et faciales [87]. De plus, certains animaux, comme les chevaux et les chiens, perçoivent nos émotions par des signaux chimiques présents dans notre odeur corporelle et produits dans des situations émotionnelles [88,89].

On décrit et réprime peut-être ses émotions différemment d'une culture à l'autre mais nous ressentons tous des émotions que nous pouvons communiquer à un animal humain ou non-humain. Leur communiquer ce que nous ressentons subjectivement, leur permet de réagir de façon adaptée. De la sorte, les réactions émotionnelles expressives et comportementales peuvent être vues comme un moyen de communication interculturel et inter-espèces.

B. Neurophysiologie de l'émotion

1. Mécanismes périphériques de l'émotion

Les réactions physiologiques et motrices associées à l'émotion impliquent nos systèmes nerveux autonome et somatique. Le comportement émotionnel est créé par deux systèmes en parallèle [45] :

- 1) *L'expression des fonctions motrices somatiques involontaires et motrices viscérales* est sous-tendue par des projections de l'hypothalamus et des centres "limbiques" du

cerveau antérieur ventro-médian vers les centres moteurs viscéraux de la formation réticulée et vers les neurones autonomes préganglionnaires du tronc cérébral et de la moelle épinière. Cela entraîne une activation des muscles lisses et des glandes.

- 2) *Le mouvement somatique volontaire* résulte des projections des zones corticales motrices du lobe frontal postérieur et du tronc cérébral vers la formation réticulée et vers les neurones moteurs somatiques du tronc cérébral et de la moelle épinière. Ceci crée des contractions musculaires et ainsi un mouvement.

2. Mécanismes centraux de l'émotion

a) Régions cérébrales de l'émotion

James Papez a proposé un ensemble de voies impliquées dans l'expression émotionnelle [90]. Il a montré que le cortex cingulaire et l'hypothalamus étaient interconnectés. Le gyrus cingulaire (avec d'autres régions corticales) se projette via le faisceau du cingulum dans l'hippocampe. L'hippocampe relaie l'information via le fornix à l'hypothalamus qui se projette à son tour vers le gyrus cingulaire. Cette boucle se nomme circuit de Papez.

Paul MacLean a ensuite révisé ce circuit en y ajoutant entre autres l'amygdale, le noyau accumbens, ainsi que différents faisceaux reliant le tronc cérébral au cerveau antérieur ou à l'hypothalamus [91].

Ces structures, avec le gyrus parahippocampique, constituent un ensemble fonctionnel appelé système limbique (Figure 5) [20,45]. Le système limbique n'est pas uniquement dédié aux émotions : il joue également un rôle dans la motivation, l'olfaction, le comportement et la mémoire.

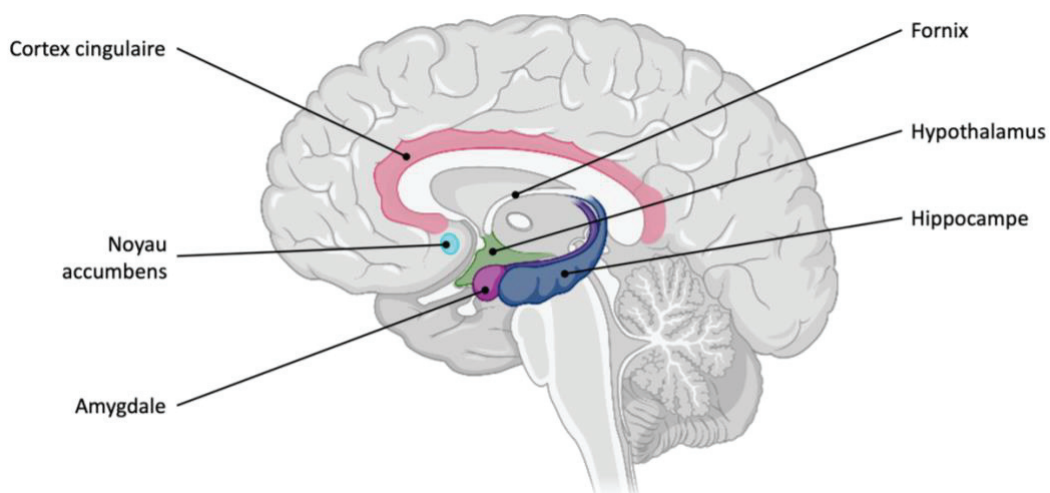


Figure 5 - Système limbique. Coupe sagittale de l'hémisphère droit mettant en évidence les systèmes neuraux traitant les émotions. (Figure créée par les auteurs sur BioRender.com)

b) Réactions cognitives et expérientielles

Les études de Joseph LeDoux suggèrent que l'activation de l'amygdale permet une évaluation émotionnelle immédiate des stimuli, qui est intégrée à d'autres types d'informations sur le stimulus [92].

La réponse émotionnelle à un stimulus dépend de notre expérience individuelle. En effet, un stimulus qui provoque du dégoût ou de la peur chez quelqu'un, peut ne pas avoir d'effet sur une autre personne. Le processus sous-jacent n'est pas encore connu.

L'amygdale semble être le point central dans ce processus du fait de ses nombreuses connexions anatomiques avec les autres régions impliquées dans l'émotion (cortex préfrontal, lobe temporal antérieur, hypothalamus, tronc cérébral, etc.). L'évaluation émotionnelle pourrait provenir de la connexion de l'amygdale avec les zones corticales orbitaires et médianes du lobe frontal qui traitent les informations en provenance de différentes modalités sensorielles. La connexion entre l'amygdale, le cortex préfrontal et les ganglions de la base servirait à sélectionner nos actions afin d'obtenir des récompenses et éviter les sanctions. L'amygdale ajouterait une touche émotionnelle à l'organisation et la planification des comportements futurs dirigés par le cortex préfrontal. L'interaction entre l'amygdale, le néocortex et les circuits sous-corticaux associés serait peut-être la source des ressentis durant l'émotion [45,93].

c) Latéralisation corticale

La théorie hémisphérique de l'émotion postule que l'hémisphère gauche est plus impliqué dans le traitement des émotions positives, tandis que l'hémisphère droit est plus impliqué dans le traitement des émotions négatives [94]. Le niveau d'intensité des émotions positives comme le bonheur est corrélé avec l'activité hémisphérique gauche, tandis que le niveau d'intensité d'une émotion négative comme le dégoût est corrélé avec l'activité hémisphérique droite [95]. On retrouve plus précisément cette asymétrie au niveau du cortex frontal. Des études ont montré que les personnes qui ont subi des dommages ou des lésions au cortex frontal gauche sont plus susceptibles de souffrir de symptômes dépressifs à la suite de ces dommages [96–100] tandis que les personnes qui ont subi des dommages ou des lésions au cortex frontal droit sont plus susceptibles de souffrir de symptômes maniaques [99–101]. De plus, on sait que les émotions positives sont associées à une motivation d'approche tandis que les émotions négatives sont associées à la motivation de retrait et il semblerait qu'une plus grande activité corticale frontale gauche est associée aux processus de motivation d'approche, alors qu'une plus grande activité corticale frontale droite est associée aux processus de motivation de retrait [102].

On notera qu'une autre théorie mentionne que l'hémisphère droit semble être plus impliqué dans la perception et l'expression des émotions que l'hémisphère gauche [95]. Effectivement, l'hémisphère droit est important dans l'expression et la détection des nuances émotionnelles de la parole [103,104]. De même, pour les expressions faciales : celles-ci sont identifiées plus facilement et plus précisément par l'hémisphère visuel droit et sont

exprimées beaucoup plus rapidement et facilement par les muscles faciaux gauches (contrôlés par l'hémisphère droit) [105].

C. Étude de l'émotion

En expérimentation, nous tentons de susciter des émotions de différentes manières : à travers l'art (film, roman, chanson), avec l'hypnose, par mouvement facial volontaire (sourire, sourcils froncés), en interagissant avec des personnes ou durant d'autres événements de la vie réelle (examen, enterrement, mariage). Dans cette section, nous allons aborder cette problématique de la mesure de ces émotions.

1. *Mesure du ressenti émotionnel*

Nous avons vu que l'émotion ressentie dépendait d'une évaluation subjective de la situation. C'est pour cette raison qu'il est important de recueillir l'état affectif ressenti directement auprès du participant ou de la participante dans le cadre d'expériences scientifiques. En effet, il existe une trop grande variabilité émotionnelle pour pouvoir proposer des stimuli créant une émotion spécifique et universelle.

Il existe différentes façons de mesurer une émotion. Le ressenti émotionnel peut simplement être demandé par un rapport oral ou écrit. Il peut aussi prendre la forme d'un questionnaire comprenant des questions ouvertes ou fermées, ou des mots à sélectionner parmi une liste d'états affectifs. Il peut également être noté quantitativement par l'utilisation d'échelles afin de noter précisément à quel point un état émotionnel est ressenti [55]. De même, il peut être demandé au participant de placer un point dans un repère orthonormé. Ce graphique en deux dimensions peut permettre de mesurer 2 échelles en même temps. Par exemple, la valence et l'éveil peuvent être mesurés de cette façon, afin de déterminer les émotions perçues à l'aide du modèle *circumplex* de Russel (Figure 4). Par ailleurs, il est possible de demander quelles émotions de base sont perçues pour en déduire *a posteriori* les émotions plus complexes ressenties à partir des combinaisons d'émotions recueillies.

On notera qu'une évaluation discrète est moins précise qu'une évaluation quantitative mais elle est plus facile à traiter par la suite. Effectivement, nous n'avons pas tous la même stratégie de notation, certains sont plus sévères, d'autres utilisent seulement les valeurs médianes et d'autres encore utilisent exclusivement les valeurs extrêmes. Lorsque des échelles sont utilisées, il est donc nécessaire de prendre en compte cette variabilité de notation avant de comparer les individus entre eux.

2. Mesure du comportement et des expressions faciales et vocales

On peut mesurer les expressions faciales à l'aide d'un électromyogramme (EMG) qui permet une mesure fine de l'activation des muscles faciaux [106]. Le comportement facial et les expressions qui en découlent peuvent également être filmés puis analysés par un algorithme ou un observateur en utilisant des méthodes comme le "Facial Action Coding System" ou le "Maximally discriminative facial motion coding system" [107]. Ces méthodes analysent la combinaison de contraction et décontraction des muscles faciaux présente pour chaque émotion.

L'entièreté du corps peut également être filmée afin d'étudier des comportements comme l'approche et le retrait par exemple [108,109]. L'enregistrement vocal permettra de reconnaître les émotions présentes tout au long d'un discours [110,111] ou d'un chant [112].

3. Mesure de l'activité du système nerveux autonome

Il est possible de mesurer l'activité physiologique d'une personne grâce à différentes technologies [55]. Dans l'étude de l'émotion, les activités périphériques que l'on mesure le plus fréquemment sont l'activité cardiaque, électrodermale, respiratoire et la température.

Concernant l'activité cardiaque la variable la plus souvent prise en compte est la fréquence cardiovasculaire [74]. D'autres mesures de la réponse cardiaque comme la variabilité de la fréquence cardiaque et la pression artérielle systolique et diastolique sont aussi couramment utilisées.

Concernant la respiration, la fréquence respiratoire est l'indice le plus souvent rapporté mais il existe de nombreuses autres variables employées pour mesurer la respiration comme la période respiratoire ou intervalle intercycle (temps entre les inspirations successives) et la profondeur de la respiration (amplitude entre l'inspiration maximale et l'expiration maximale) [113].

L'activité électrodermale correspond aux variations électriques induites par l'activité de sudation des glandes sudoripares présentes dans notre peau. La variable d'activité électrodermale la plus souvent utilisée est le niveau de conductance cutanée, suivie de l'amplitude de cette réponse et du taux de réponse cutanée [74]. En réaction à un stimulus, des pics d'activité électrodermale peuvent parfois apparaître, on appelle cela la réponse de conductance de la peau (SCR, pour Skin Conductance Response en anglais). La SCR qui apparaît à un moment qui n'est pas associé à un événement particulier est nommée réponse conductance de la peau non spécifique (NSSCR).

4. Mesures de l'activité du système nerveux central

Il existe différents moyens d'enregistrer l'activité cérébrale comme l'imagerie par résonance magnétique (IRM), l'électroencéphalographie (EEG) et la tomographie par émission de positons (PET). L'IRM et l'EEG ont l'avantage d'être non invasifs.

a) IRM

L'Imagerie par Résonance Magnétique (IRM) permet d'obtenir des images en 3 dimensions de l'intérieur du corps. La technique rend possible la visualisation de notre organisme en se déplaçant dans les différents plans de coupes anatomiques : transversal, coronal et sagittal (Figure 6).

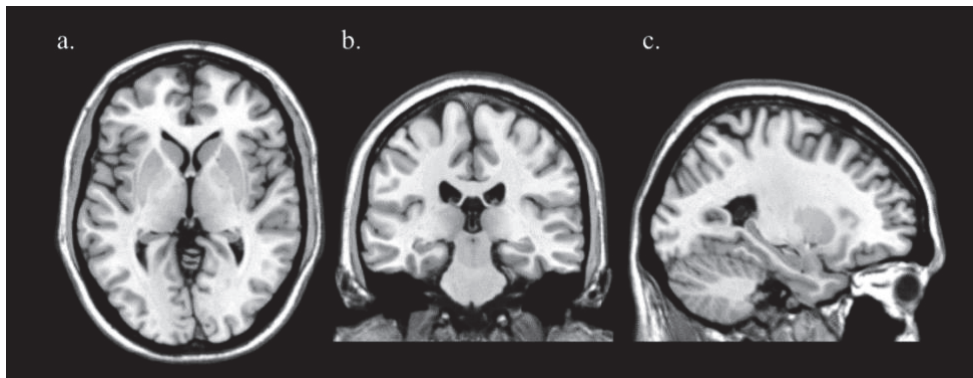


Figure 6 - Image type IRM cérébrale. a) Plan transversal ou axial b) Plan coronal c) Plan sagittal

L'IRM est la technologie non invasive ayant la meilleure résolution spatiale. On peut obtenir des coupes du cerveau entier avec une résolution spatiale allant en moyenne de 2 à 4 mm et une résolution temporelle entre 1 à 2 s. Sa résolution temporelle est assez faible, car les différentes coupes du cerveau sont enregistrées de façon asynchrone les unes à la suite des autres. Ces résolutions spatiale et temporelle sont dépendantes, augmenter l'une diminue l'autre. L'intensité du champ magnétique a également un impact sur la résolution temporelle, plus celui est fort, plus la résolution sera élevée [114].

L'imagerie par résonance magnétique fonctionnelle cérébrale (IRMf) mesure et localise l'activité cérébrale en détectant les changements locaux de flux sanguin. Les molécules d'hémoglobine contenues dans nos vaisseaux sanguins se nomment oxyhémoglobines lorsqu'elles sont oxygénées et désoxyhémoglobines lorsqu'elles sont désoxygénées. L'activité neurale provoque deux phénomènes vasculaires en parallèle : une augmentation de la consommation d'oxygène par les neurones, ce qui réduit la quantité d'oxyhémoglobine dans les vaisseaux sanguins de cette zone, et l'augmentation du flux sanguin, ce qui augmente la quantité d'oxyhémoglobine dans les vaisseaux sanguins de cette zone cérébrale. L'augmentation du débit sanguin est plus importante que la consommation d'oxygène par les neurones, cela entraîne ainsi une réduction relative de désoxyhémoglobine dans les veines avoisinant les neurones actifs (Figure 7). La concentration de désoxyhémoglobine peut être détectée par l'IRM. Par la

libération d'oxygène, la désoxyhémoglobine contient deux électrons libres. Ces électrons créent une modification du champ magnétique local généré par le scanner IRM. Cette modification du champ magnétique est détectée et enregistrée sous forme d'un signal appelé BOLD pour Blood Oxygen Level Dependent [115,116].

Les images d'IRMf sont bruitées et contiennent des artefacts liés aux mouvements de l'individu et des distorsions spatiales. Ainsi il est nécessaire d'enregistrer plusieurs fois le signal IRM associé à un même stimulus. Le signal doit également être prétraité : il est généralement lissé et filtré pour réduire le bruit et des corrections sont appliquées pour limiter les artefacts.

Il est difficile d'interpréter avec précision les images fonctionnelles (ou scans) d'un seul individu parce que l'amplitude du signal et la forme du cerveau sont propres à chacun. Pour neutraliser ces différences il est courant de normaliser le signal et de réaliser une normalisation spatiale dans un modèle de cerveau.

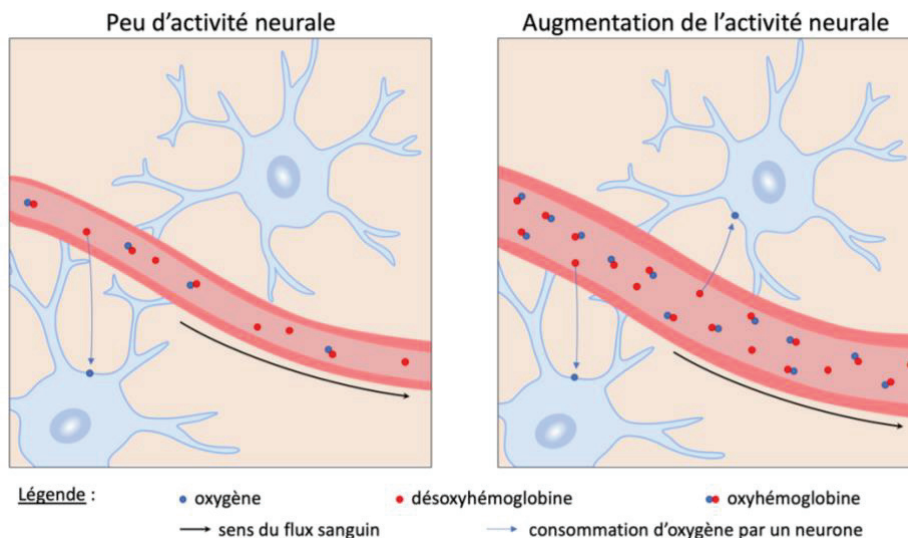


Figure 7 - Effet BOLD. L'augmentation de l'activité neurale entraîne une augmentation du flux sanguin et une diminution de la proportion de désoxyhémoglobine (hémoglobine non oxygénée) détectable par l'IRM.

b) EEG

L'électroencéphalographie est une technique permettant de mesurer l'activité électrique générée par le cerveau. Les neurones pyramidaux des couches II à VI du cortex émettent une activité électrique. Celle-ci peut être captée par des électrodes placées sur le cuir chevelu puis amplifiée, filtrée, et numérisée pour enfin être enregistrée sous forme de tracé ou électroencéphalogramme.

Cette technologie a une assez faible résolution spatiale car le liquide céphalo-rachidien, les méninges, le crâne et la peau dévient le champ électrique de plusieurs millimètres par rapport à sa source. Elle ne permet pas l'acquisition d'une image tridimensionnelle comme l'IRMf et le PETscan. Elle possède cependant une excellente résolution temporelle (de l'ordre de la milliseconde).

L'activité cérébrale mesurée par l'électroencéphalographie provient des potentiels postsynaptiques d'excitation et d'inhibition qui arrivent au niveau des nombreuses dendrites des neurones pyramidaux. Ces mouvements de particules chargées (ions Na^+ , K^+ , Ca^{++} , Cl^-) dans les milieux intra et extracellulaires peuvent être captés par des électrodes généralement recouvertes de chlorure d'argent. L'activité enregistrée par ces capteurs correspond à la somme des activités simultanées des neurones pyramidaux, il faut donc un grand groupe de neurones actifs pour obtenir une activité électrique au niveau du scalp (cuir chevelu). Cette activité électrique cérébrale est présente en permanence et spontanément. L'enregistrement de cette activité de base se nomme EEG continu. Une stimulation externe (environnementale) ou interne (cognitive) peut avoir deux types d'effet sur l'activité cérébrale spontanée :

- *Un potentiel évoqué (ou potentiel lié à l'évènement)*. Une réponse spécifique sous forme de fluctuations importantes de tensions est créée par le stimulus, qui peut être associée à une émotion, une pensée, un souvenir, une perception visuelle d'image, un mot ou une stimulation tactile.
- *Une synchronisation ou désynchronisation liée à l'évènement*. L'état de fonctionnement des neurones actifs est modifié par un stimulus tel l'ouverture des yeux, un son ou une stimulation nociceptive.

Le signal étant très faible, il est nécessaire d'enregistrer plusieurs fois le signal EEG associé à une stimulation donnée durant l'expérience afin de moyenniser le signal et de dissocier la réponse au stimulus de l'activité de fond. Les données générées par l'électroencéphalographie peuvent être difficiles à interpréter de par leur volume important et du fait que les schémas d'ondes cérébrales sont propres à chaque individu [117–121].

En conclusion, à l'aide des différentes technologies à notre disposition, nous allons pouvoir mesurer l'émotion sous différents angles (ressenti subjectif, activité périphérique, activité centrale) et étudier différents paramètres (comme la spatialité ou la temporalité pour l'activité cérébrale). Un invariant sera tout de même présent dans les différentes études présentées : les stimuli utilisés pour induire des émotions seront des odeurs. Ainsi, nous proposons, ici, d'étudier l'olfaction, un sens qui semble justement avoir un lien tout particulier avec les émotions.

II. Les odeurs comme vecteur d'émotions

Nous interagissons avec l'environnement grâce à nos systèmes sensoriels. Parmi ces derniers, le sens de l'odorat, nous permet de détecter, d'identifier et de discriminer les molécules odorantes présentes dans l'air.

Cette section a pour objectif de fournir des éléments d'information sur le rôle et le fonctionnement de l'olfaction chez l'humain, sur son rapport privilégié avec les émotions et sur sa complexité en termes de diversité inter et intra-individuelle.

A. L'olfaction : un sens sous-estimé

Le sens de l'odorat est un sens méconnu, souvent sous-estimé et négligé. Il est d'ailleurs considéré pour beaucoup comme le moins aiguë et utile de nos sens. De plus, il est souvent confondu avec d'autres modalités sensorielles comme la gustation (sucré, salé, amer, acide, umami) et la sensibilité trigéminal (piquant, irritant, froid, chaud).

Pour comprendre l'importance de l'olfaction, il est intéressant de prendre en compte le témoignage de ceux qui en sont dépourvus.

1. *Pertes olfactives et qualité de vie : témoignages*

Il existe différentes causes possibles de troubles de l'odorat.

Les troubles de l'odorat - dits quantitatifs - sont de deux types : l'anosmie (perte totale de l'odorat) et l'hyposmie (perte partielle de l'odorat) [122]. La perte olfactive peut être congénitale (de naissance) [123,124], due au vieillissement (presbyosmie) [125–127], survenir à la suite d'un accident (lésions post traumatiques) [128] ou à une maladie infectieuse (ex. virus) [129], rhinosinusienne [130] ou neuro-dégénérative (Parkinson, Alzheimer) [131]. L'hyposmie implique aussi bien une faible sensibilité, c'est-à-dire un seuil élevé de détection des odeurs, que des difficultés à discriminer des odeurs ou encore une identification des odeurs erronée ou inaccessible. En 2015, en France, entre 10 et 15% des jeunes adultes, et entre 30 et 40% des personnes âgées présentaient des troubles olfactifs [132]. Ces chiffres ont considérablement augmenté ces trois dernières années avec la pandémie de Covid-19 [133].

La perte olfactive est un handicap qui a des répercussions sur la vie des individus. En effet, certains aspects de la qualité de vie comme les préférences et habitudes alimentaires sont altérés chez les personnes atteintes de troubles olfactifs. Souvent, les patients se plaignent d'une diminution de l'appétit, de rencontrer des difficultés à cuisiner et à détecter des aliments avariés ou brûlants [130]. Ils se sentent également en insécurité, car pour beaucoup ils ne peuvent pas détecter l'odeur du feu, de la fumée ou du gaz [134]. De plus, ces pertes olfactives influencent leur hygiène personnelle. Leur odeur corporelle ou celle de leurs enfants est d'ailleurs une de leurs principales inquiétudes [135]. Ils ont peur, par exemple, de sentir mauvais, de mettre trop de parfum ou de ne pas être informés qu'il faut changer la couche de leur bébé. La perte de l'odorat a également d'autres effets sociaux

comme une diminution de l'appétit sexuel [134,135]. Prises dans leur ensemble, ces données pointent vers un impact significatif sur le bien-être des personnes touchées : un tiers des personnes souffrant de troubles olfactifs sont sujettes à la dépression [136,137].



CONTRIBUTION SCIENTIFIQUE **Perte de l'odorat avec la COVID-19**

J'ai participé à deux études au sujet de la perte de l'odorat liée à la COVID-19, la première analysait l'évolution de la pandémie à travers le nombre de personnes auto-déclarées anosmiques ou hyposmiques. La seconde étude s'intéressait à la durée de la perte de l'odorat et l'impact sur la qualité de vie des patients. Les articles publiés suites à ces études sont disponibles en [Annexe 2](#) et [Annexe 3](#).

2. Principales fonctions de l'odorat

L'olfaction a trois fonctions essentielles. La première est celle de la **prévention d'un danger** (menace bactérienne/fermentation, fuite de gaz, fumée toxique, produit chimique) souvent signalé par la présence d'une odeur que notre cerveau interprète comme « dangereuse » ou « désagréable » [138,139].

Sa deuxième fonction principale est à mettre en lien avec l'**alimentation** : l'odorat régule notre appétit en nous indiquant quand, quoi et en quelle quantité manger. Il nous permet également de localiser les aliments, de déterminer leur comestibilité, de les cuisiner et de les savourer [139,140].

Sa troisième fonction concerne la **communication sociale** : l'olfaction joue entre autres un rôle important dans la contagion émotionnelle [141] et est impliqué dans l'attirance physique entre deux individus [142,143]. Elle nous aide également à maîtriser ou contrôler notre odeur corporelle [139,144].

B. De la molécule chimique à la perception de l'odeur

1. Odeur et odorant

Une odeur est une perception qui résulte de l'interaction entre une molécule ou un mélange de molécules émises par une source dans l'environnement et nos récepteurs olfactifs (chimiorécepteurs) présents dans notre cavité nasale. Toutes les molécules ne sont pas odorantes ; en effet, il faut qu'elles remplissent un certain nombre de propriétés pour qu'elles puissent être aéroportées dans notre cavité nasale et interagir avec les récepteurs protéiques situés sur les neurones olfactifs qui tapissent le plafond de la cavité nasale. Les critères qui font qu'une molécule chimique peut être un odorant sont les suivants :

- 1) Elle doit être **volatile** afin qu'elle puisse se propager dans l'air ambiant et entrer dans notre nez lorsque nous inspirons. Pour cela, il ne faut pas qu'elle soit trop légère ni trop lourde (environ entre 26 à 300g/mol) [145].

- 2) Elle doit se trouver en assez **grande concentration** dans l'air. En effet, bien que cela dépende de notre propre sensibilité, il faut au minimum 10^{12} molécules par litre d'air pour pouvoir détecter une molécule odorante. [146]
- 3) Elle doit être **soluble** dans l'eau afin de se dissoudre dans le mucus qui tapisse l'épithélium olfactif. [146,147]
- 4) Elle doit être porteuse d'un **groupe osmophile**, c'est-à-dire d'un groupe caractéristique d'atomes pouvant interagir avec les récepteurs et déclencher un message nerveux. [146,148]

Ces critères sont nécessaires mais pas suffisants car ils n'entraînent pas forcément la perception d'une odeur. En effet, d'autres caractéristiques peuvent avoir une influence sur l'interaction ou non avec les récepteurs comme sa polarité, sa structure, sa chiralité et son isomérisation [146,149].

De plus, deux molécules proches chimiquement n'auront pas forcément une odeur semblable. Prenons, par exemple, deux énantiomères c'est-à-dire deux molécules dont seule la conformation dans l'espace change comme le (1S,4R)-cis-*p*-menthan-8-thiol-3-ones et le (1R,4S)-cis-*p*-menthan-8-thiol-3-ones. Le premier sent une odeur fruitée comme la feuille de cassis et le fruit de la passion alors que le second dégage une désagréable odeur de caoutchouc et de brûlé [150]. À l'inverse deux molécules très différentes chimiquement peuvent être perçues comme semblables : il existe de nombreuses molécules qui sentent la rose comme le 2-phényléthanol, le nérol et le géraniol.



MÉDIATION SCIENTIFIQUE

Jeux vidéo avec les récepteurs olfactifs

J'ai participé à la Scientific Game Jam dont l'objectif est de créer, en 48h, un jeu vidéo autour de la thématique de recherche d'un-e doctorant-e. Mon équipe et moi avons conçu un jeu d'arcade nommé « Le lac des cils » qui se déroule au niveau des récepteurs olfactifs.

Le jeu se déroule dans la cavité nasale d'un personnage fictif qui « sniffe » un parfum. L'objectif est d'emmener les molécules jusqu'au cil olfactif en accrochant le bon groupe osmophile aux bons neurones (système clé-serrure). Si le joueur a un assez bon score, le personnage sent l'odeur et on aperçoit sa réaction. Ainsi le joueur peut découvrir :

- Qu'une odeur contient différentes molécules qui activent les récepteurs présents dans le mucus de l'épithélium olfactif.
- Qu'un neurone comporte les mêmes récepteurs olfactifs sur l'ensemble de ces cils.
- Qu'une même molécule peut se fixer à différents types de récepteur olfactif et qu'un même neurone peut être activé par différentes molécules.
- Les différents rôles de l'odorat à travers les réactions des personnages face à l'odeur.
- Les différences entre sujets sains et malades (covid-19 et polypes).

Le prototype du jeu conçu est jouable sur itch.io : <https://bit.ly/3A8FTYE>

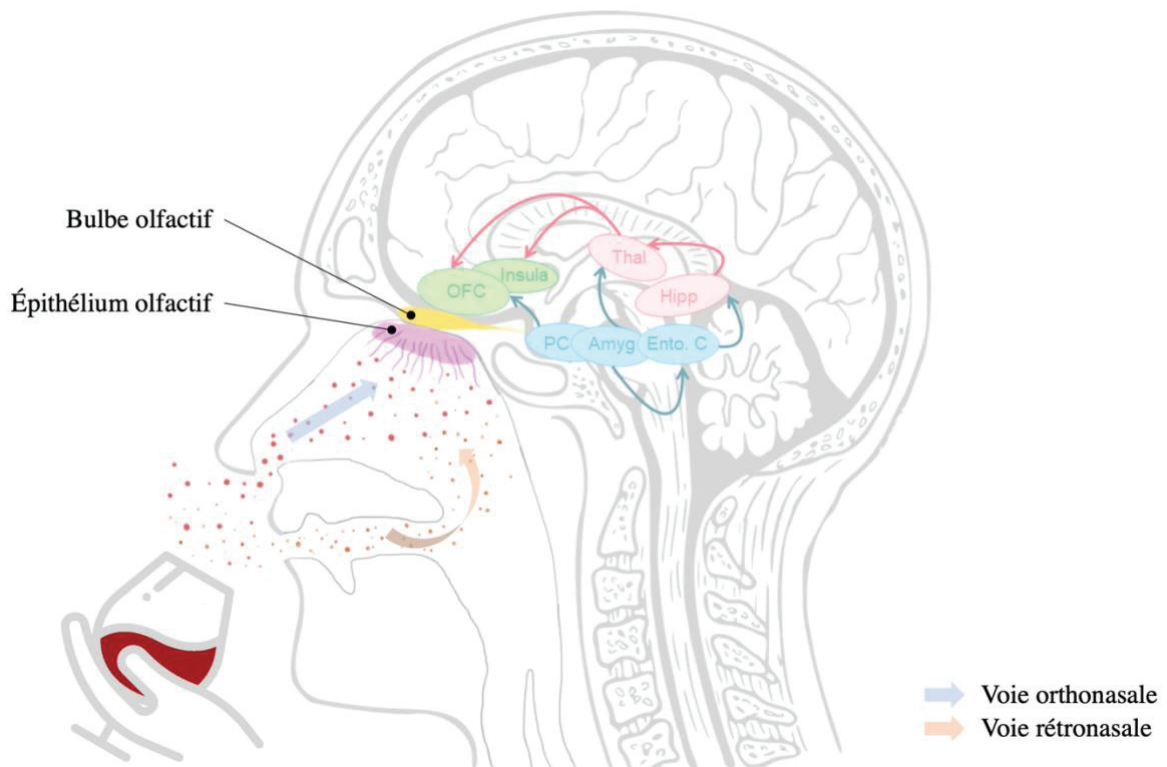


Figure 8- Fonctionnement de l'olfaction. En bleu, le cortex olfactif primaire avec l'amygdale (Amyg), le cortex entorhinal (Ento. C) et le cortex piriforme (PC). En vert, le cortex olfactif secondaire constitué de l'insula et du cortex orbitofrontal (OFC). En rose l'hippocampe (Hip) et le thalamus (Thal).

2. Mécanismes périphériques de l'olfaction

Il existe deux voies olfactives différentes pour stimuler les récepteurs olfactifs (Figure 8). Lorsque l'on sent en inspirant par les narines, on utilise la voie orthonasale. Lorsque l'on sent des arômes par la bouche, les odeurs remontent dans la fosse nasale par l'arrière du palais, c'est la voie rétronasale. Ces voies transportent les molécules jusqu'à la cavité nasale, et c'est ici que commence l'odorat.

L'odorat débute plus particulièrement au niveau de la muqueuse olfactive. Cette dernière mesure environ 5 cm² et est positionnée dans le cornet supérieur et forme quelques îlots dans le cornet médian et sur le septum nasal. C'est à ce niveau qu'a lieu la transformation du signal chimique en message nerveux (transduction olfactive). La muqueuse olfactive est composée du mucus olfactif et de l'épithélium olfactif qui contient entre autres les neurones olfactifs (Figure 9).

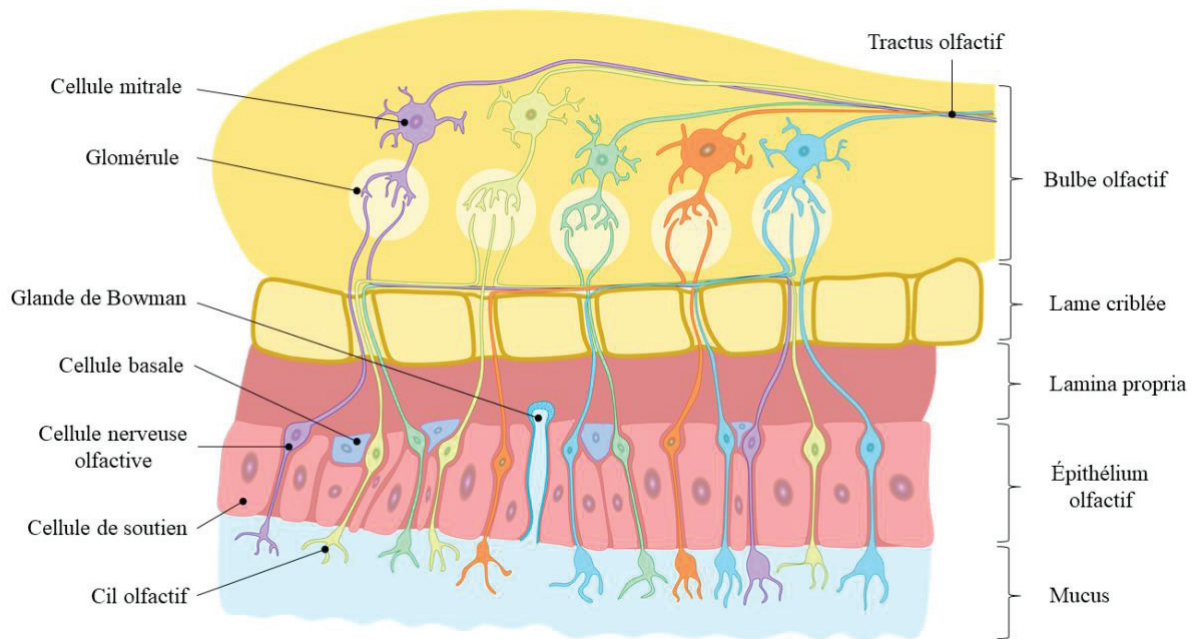


Figure 9 – Représentation d'une coupe d'une muqueuse olfactive et d'un bulbe olfactif. Les neurones olfactifs sont représentés en 5 couleurs différentes : les neurones d'une même couleur expriment le même type de récepteurs et se projettent vers le même glomérule. Lorsque des molécules sont aéroportées dans la cavité nasale, au niveau du mucus, elles interagissent avec les cils des cellules nerveuses olfactives et créent un message nerveux arrivant aux cellules mitrales du bulbe olfactif, qui transmettent l'information au reste du cerveau via le tractus olfactif.

a) Le mucus

C'est dans le mucus que se trouvent les récepteurs olfactifs, ceux-ci sont donc en contact direct avec l'extérieur. Le mucus, sécrété par les glandes de Bowman, recouvre et protège l'épithélium olfactif. Il est constitué d'une substance aqueuse dans laquelle les odorants se solubilisent pour atteindre les cellules olfactives de l'épithélium soit par diffusion soit par fixation sur une protéine de transport appelée OBP (*Odorant Binding Protein*). Comme illustré Figure 10, l'OBP est une lipocaline (protéine de transport) qui a pour rôle de concentrer près des récepteurs les molécules odorantes solubilisées dans différentes régions de la cavité nasale.

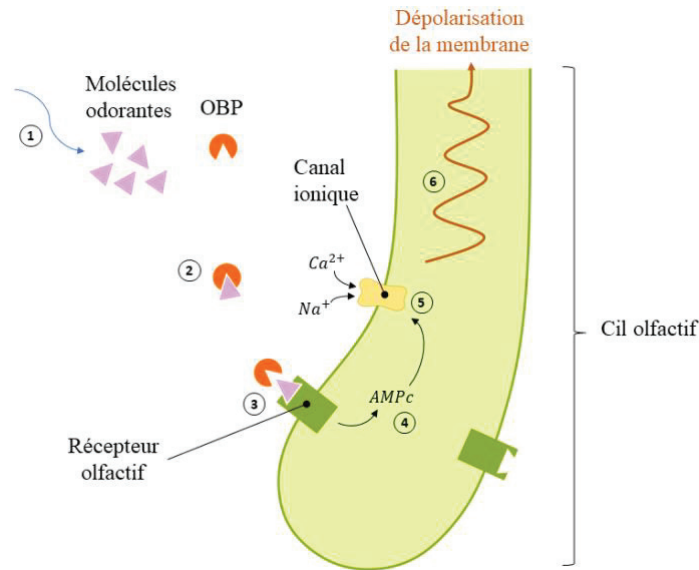


Figure 10 – Transduction olfactive. 1. Les molécules odorantes arrivent dans le mucus. 2. La molécule odorante se fixe à l'OBP qui la transporte vers le cil olfactif. 3. La molécule interagit avec un récepteur du cil olfactif et 4. Cette interaction produit de l'AMPc. 5. Cela déclenche l'ouverture de canaux ioniques et 6. Cela crée une dépolarisation de la membrane transmise en direction du bulbe olfactif.

b) L'épithélium olfactif

Pour faire face aux agressions extérieures, l'épithélium olfactif se compose de 3 types de cellules (Figure 9) : les neurones olfactifs qui comportent les récepteurs olfactifs, les cellules de soutien qui structurent la muqueuse olfactive et entretiennent les neurones olfactifs, et les cellules basales qui servent à renouveler régulièrement la muqueuse. Ces dernières peuvent se différencier en neurones olfactifs [151]. De cette façon, les cellules olfactives sont capables de neurogenèse : elles peuvent renouveler régulièrement les cellules détériorées. Cette génération cellulaire n'étant pas synchronisée entre tous les neurones olfactifs, la durée de vie d'un neurone olfactif chez l'humain n'est pas connue [152]. Elle semble être d'environ 30 jours chez les mammifères [153].

Les cellules nerveuses olfactives sont des neurones bipolaires : leur axone part directement dans le cerveau au niveau du bulbe olfactif tandis que les 8 à 20 cils de leurs dendrites s'étendent jusqu'à la surface de l'épithélium où ils baignent dans le mucus. Les récepteurs olfactifs sont enchâssés dans la membrane des cils olfactifs (Figure 10).

Nous disposons d'une grande diversité de récepteurs olfactifs, nous permettant de détecter et de distinguer un large spectre de molécules odorantes. Une molécule odorante peut activer un ou plusieurs des 400 types de récepteurs différents dans notre épithélium olfactif, et en retour un type de récepteur peut réagir à une ou plusieurs molécules odorantes. C'est un codage combinatoire pratiquement infini à tel point que l'humain aurait la capacité de discriminer plus de 1000 milliards d'odeurs [154].

Chaque neurone olfactif exprime toujours un seul type de récepteur. Comme illustré dans la Figure 10, lorsqu'une molécule se fixe sur un récepteur, l'étape de transduction olfactive a lieu : cela induit la formation d'AMPc (adénosine monophosphate cyclique) ou d' IP_3 (inositol triphosphate) qui déclenche l'ouverture des canaux ioniques de la membrane. Des échanges d'ions Na^+ , Ca^{2+} ont lieu entre l'intérieur et l'extérieur de la cellule, pour

conduire à une dépolarisation de la membrane et créer un potentiel d'action (PA) [45,155]. Ce PA, ou message nerveux, sera alors transmis au bulbe olfactif par les axones des neurones olfactifs qui traversent la boîte crânienne par la lame criblée de l'ethmoïde (Figure 10).



MÉDIATION SCIENTIFIQUE **Mini-série d'animation sur l'olfaction**

Cette partie est illustrée dans une fiction animée qui explique le trajet d'une molécule odorante jusqu'au récepteur olfactif. Cette vidéo est le premier épisode d'une mini-série d'animation réalisée en collaboration avec des chercheurs et des artistes. Ce projet a été sélectionné dans le cadre de l'appel à projet 2021 du GDRO3.

Deux autres épisodes ont également été conçus, l'un détaille les trois sens chimiques (gustation, olfaction et trigéminal) et le dernier traite de l'anosmie.

Lien : <https://bit.ly/3p8CyCA>

3. Mécanismes centraux de l'olfaction

Le bulbe olfactif traite le message olfactif en provenance des neurones récepteurs de la cavité nasale et relaie l'information au cortex olfactif primaire qui projette à son tour vers d'autres aires cérébrales dites secondaires.

a) Le bulbe olfactif

Toutes les cellules olfactives qui expriment le même type de récepteur font converger leurs axones vers une même région du bulbe olfactif appelée glomérule (Figure 9) où elles établissent un contact synaptique avec les dendrites des cellules mitrales et celles des cellules à panache. Certains interneurons font également l'intermédiaire entre ces cellules : les cellules granulaires relient les cellules mitrales entre elles et les cellules périglomérulaires modulent la transmission de l'information entre les neurones récepteurs olfactifs et les cellules mitrales au niveau des glomérules [156].

b) Le cortex olfactif primaire et secondaire

Les axones des cellules mitrales et des cellules à panache se projettent dans le cortex olfactif primaire incluant le noyau olfactif antérieur, le tubercule olfactif, ainsi que les cortex piriforme, périamygdalien et entorhinal. Les neurones de ces régions se projettent ensuite sur l'hippocampe, l'amygdale, le thalamus, et également sur le cortex orbitofrontal et l'insula (Figure 8). L'amygdale transmet aussi l'information à l'hypothalamus et au mésencéphale [157,158].

Par rapport aux autres sens, la voie empruntée par le signal olfactif est plus directe : elle ne passe pas obligatoirement par le thalamus avant d'atteindre les aires corticales [159] et

les régions cérébrales fortement impliquées dans la mémoire (cortex entorhinal, hippocampe) et dans les émotions (amygdale) [160]. Il semble exister un lien anatomique privilégié entre l'olfaction, la mémoire [161] et les émotions [162,163].

C. Le pouvoir émotionnel des odeurs

Nous allons aborder ce lien entre olfaction et émotions, tout d'abord, sous un angle linguistique en nous intéressant à la charge affective des mots liés aux goûts et aux odeurs dans le langage courant anglais ou français. Nous nous intéresserons également à notre façon de décrire les odeurs. Ensuite, nous aborderons les connaissances présentes dans la littérature concernant la manifestation des émotions induites par un stimulus olfactif dans le système nerveux central et périphérique.

1. Langage

a) Langage courant

Le vocabulaire émotionnel est riche en termes olfactifs et gustatifs. Bien que le langage en lien avec les sens utilise principalement des caractéristiques visuelles [164], les phrases à valence émotionnelle sont davantage chargées en mots provenant du champ lexical de l'odorat et du goût par rapport aux autres sens [165]. Ainsi des adjectifs de l'olfaction et la gustation permettent d'ajouter une dimension émotionnelle à certaines phrases par exemple : « un bain parfumé » ou « un délicieux moment ».

Réciproquement, le vocabulaire olfactif et gustatif est plus chargé émotionnellement que celui des autres sens [165]. En effet, certains mots évoquant des perceptions visuelles, auditives et tactiles, comportent une teinte émotionnelle comme « beau » et « bruyant » mais la majorité d'entre eux en sont dépourvus (ex. « brun », « écho », « lisse »). Les termes olfactifs et gustatifs, en revanche, ont en moyenne une plus forte valence émotionnelle (ex. « saveur », « âcre »).

De plus, même les verbes de perception de base tels que "sentir" et "goûter" présentent une composante émotionnelle [166,167]. En français, on le remarque particulièrement dans des expressions comme « ne pas le sentir » ou « avoir du nez ».

b) Description d'odeur

Une odeur peut être décrite par différentes dimensions comme :

- 1) Sa qualité, son identité : ex. framboise.
- 2) Son intensité : si l'odeur perçue de manière faible ou forte. Cette perception dépend, entre autres, de la concentration de la molécule dans l'environnement.
- 3) Sa familiarité : si la personne ne connaît pas l'odeur ou si elle y est habituée.
- 4) Sa valeur hédonique : caractère plaisant/déplaisant (odeur agréable vs désagréable).
- 5) Sa source.
- 6) La comestibilité (ou non) de sa source.

Pourtant la population non experte en olfaction éprouve des difficultés à décrire les odeurs. En effet, une personne peut déclarer sentir une odeur mais il est fort probable qu'elle l'identifie à tort ou même qu'elle n'arrive pas à l'identifier [168].

Nous nommons les odeurs par rapport à leur source (ex. rose, menthe), cependant, les individus éprouvent des difficultés à associer une odeur à un nom même si cette odeur leur est très familière. Il existe d'ailleurs une expression pour désigner le fait de ne pas être capable de nommer une odeur connue : « sur le bout du nez ». Contrairement à l'état « sur le bout de la langue », donner un indice sur le nom comme la première lettre ou le nombre de syllabes n'aide pas toujours la personne à trouver le nom de l'odeur qu'elle a « sur le bout du nez » [169].

Décrire une odeur par un ensemble de caractéristiques, comme nous utilisons les couleurs et les formes pour écrire une image, n'est pas habituel dans la vie d'une personne non experte. Il existe d'ailleurs peu de mots dans le français et l'anglais courant qui identifient des propriétés précises des odeurs. Les mots couramment utilisés comme « puant » décrivent l'affect et les mots comme « réglisse » identifient la source de l'odeur et non une qualité de l'odeur [168].

Contrairement aux goûts et aux couleurs, nous n'apprenons pas à décrire les odeurs ; les odeurs sont ainsi construites au sein de notre expérience subjective. De cette façon, la signification d'une odeur pour l'individu n'est rattachée à aucune norme sociale. L'expérience olfactive est davantage conçue comme une expérience personnelle (individuelle) que collective (partagée au sein d'un groupe) [169–171].

Cela se retrouve dans les descriptions d'odeurs qui incluent assez fréquemment une implication personnelle du locuteur par l'utilisation de pronom personnel (moi, je, nous) et d'adjectif possessif (mon, ma, notre...) ou l'évocation de souvenirs personnels, par exemple : « Je ne connais pas cette odeur, mais j'aime bien. Je trouve qu'elle détend, elle est relaxante. Ça me fait penser à la nature. » ou « Pour moi, l'odeur est plutôt... pénible. Ça m'agresse le nez. » [170].

Une analyse de corpus [172] en français, a identifié différentes propriétés utilisées pour construire la dénomination d'une odeur. Les principaux axes de classification étaient l'effet que l'odeur procure à la personne (pour 41% des réponses étudiées) et la source de l'odeur (27%). Les réponses restantes concernaient d'autres propriétés comme l'intensité (7%) et la mémoire (5%). La mémoire n'est pas la composante la plus fréquente mais il est intéressant de constater que les souvenirs évoqués par les odeurs sont particulièrement chargés émotionnellement par rapport aux souvenirs en lien avec d'autres modalités sensorielles [165,171].

La composante la plus fréquente est celle de la description des effets de l'odeur sur celui ou celle qui la perçoit [171,172]. Cette description des manifestations d'une sensation indissociable de la personne sentant l'odeur est associée au plaisir et déplaisir ressenti par le sujet [171]. Le jugement hédonique est d'ailleurs souvent considéré comme la principale caractéristique de la description des odeurs [173,174]. En effet, il est un paramètre déterminant pour créer une émotion guidant les comportements d'approche ou d'évitement et permet ainsi à l'olfaction de remplir ses fonctions (cf. II.A.2.Principales fonctions de l'odorat).

Pour résumer, nous n'avons pas tous le même vécu olfactif, nous n'avons pas appris de lexique spécifique pour la description d'odeur et même lorsque nous connaissons une odeur, nous avons du mal à retrouver l'objet-source associé. Par conséquent, nous avons une perception subjective des odeurs qui nous amène à les décrire de façon émotionnelle.



CONTRIBUTION SCIENTIFIQUE **Description subjective d'odeurs**

En collaboration avec une autre doctorante, Marylou Mantel, j'ai participé à une étude d'analyse de descriptions spontanées d'odeurs. Ma contribution a consisté à inclure dans cette publication une analyse data mining de type *Découverte de sous-groupes* dans le but de caractériser les différentes notions et concepts abordées par des populations différentes (sexe, langue, fréquence d'expositions) en réponse à des odeurs. J'ai également intégré à la publication une analyse en *Clustering* avec une description des clusters par recherche d'itemsets fermés fréquents pour expliciter les différentes perceptions d'une même odeur. L'article résultant est consultable en [Annexe 1](#).

2. Odeurs, émotions et système nerveux

a) Système nerveux central

La dimension hédonique étant reconnue comme la dimension la plus saillante de l'olfaction, la localisation des représentations hédoniques plaisantes et déplaisantes dans le cerveau a été étudiée par plusieurs groupes dans la communauté scientifique.

Des études montrent que les odeurs agréables et désagréables induisent une plus grande réponse neuronale dans l'amygdale [175] et le cortex piriforme antérieur [176] par rapport aux odeurs neutres. Le cortex olfactif secondaire, composé de l'insula et du cortex orbitofrontal, est également impliqué dans la perception hédonique des odeurs [177–179]. En effet, cette région représente – entre autres – la valeur de la récompense alimentaire [179–181].

La zone piriforme-amygdale avec l'insula ventrale est davantage activée par les odeurs désagréables que par les odeurs agréables [178]. Cette tendance est observable particulièrement au niveau de l'amygdale [176,178,182], de l'insula gauche [178,183] et du cortex piriforme frontal [184].

Certaines études montrent également une latéralisation de ce pattern d'activité dans la partie gauche du cortex piriforme antérieur [183,185]. Une publication montre également que cette réponse élevée du cortex piriforme antérieur ou frontal aux odeurs désagréables diminue avec le temps alors que l'activité pour les odeurs agréables reste constante. Il y aurait donc un effet de l'habituation dans cette région, effet qui dépendrait de la valence du stimulus [176].

Le cortex piriforme postérieur serait plutôt impliqué dans la représentation de la qualité perceptive de l'odeur. En effet, celui-ci semble activé bilatéralement par toutes les odeurs, indépendamment de la valence [176,186], et il s'activerait davantage pour les odeurs intenses que faibles [177,187].

D'autres structures sont également impliquées dans le traitement de l'agrément comme le tubercule olfactif [188], le gyrus cingulaire antérieur et l'aire tegmentale ventrale [189].

Cependant ces résultats ne sont pas consensuels puisque quelques études montrent que l'activité au sein de l'amygdale et du cortex piriforme ne seraient pas associée à la valence mais plutôt à l'intensité [177,187,190]. D'autres encore suggèrent que l'amygdale ne code ni l'intensité ni la valence en soi, mais une combinaison des deux dimensions qui reflète la valeur émotionnelle globale d'un stimulus [191].

b) Système nerveux périphérique

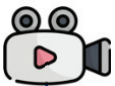
Le processus olfactif affectif a également été étudié au niveau du système nerveux périphérique et essentiellement à travers des paramètres comme la respiration, le rythme cardiaque, la température de la peau et la conductance cutanée.

Concernant l'activité respiratoire, les odeurs agréables augmentent le volume d'air inspiré, diminuent la fréquence respiratoire : de la sorte, elles entraînent une respiration lente et profonde. À l'opposé, les odeurs désagréables diminuent le volume d'air inspiré et augmentent la fréquence respiratoire : elles créent ainsi une respiration rapide et superficielle [192].

Pour la température cutanée, la plupart des études ne trouvent pas de différenciation entre les odeurs plaisantes et déplaisantes [193]. Une étude rapporte cependant une température cutanée plus haute pour l'odeur désagréable dans un premier temps puis plus basse après quelques secondes [194].

La fréquence cardiaque semble corrélée au caractère hédonique : plus une odeur est désagréable (resp. agréable) plus le rythme cardiaque augmente (resp. diminue) [193–196].

En ce qui concerne l'activité électrodermale, les odeurs désagréables se caractérisent par une réponse électrodermale plus longue [193,197] et plus élevée en termes d'amplitude que les odeurs agréables [194,196,198–200]. Un second phénomène est observé à propos de l'amplitude du pic de conductance cutanée : la réponse électrodermale augmente avec l'intensité et l'éveil émotionnel de l'odeur [178,195].



MÉDIATION SCIENTIFIQUE Vidéo de vulgarisation scientifique

Concernant le lien entre l'olfaction et les émotions, j'ai collaboré avec l'association « L'exploratoire » pour réaliser une courte vidéo de vulgarisation sur le sujet. Cette vidéo est disponible sur la chaîne youtube 99sec.

Lien : <https://bit.ly/3P535LO>

D. La variabilité individuelle dans la perception olfactive

Une même odeur peut être perçue différemment par des individus distincts mais peut également être perçue différemment chez un même individu à différents moments de sa journée, de sa vie. Dans cette section nous allons parcourir ces facteurs de variabilité inter et intra-individuelle en matière d'odorat.

1. Génétique

Nous avons vu que nos 400 récepteurs olfactifs nous permettent de discriminer un grand nombre d'odeurs [201]. Ces récepteurs sont des protéines et la présence de ces dernières sur la membrane des neurones olfactifs dépend de notre génome. Trois à 5% de nos gènes codent pour des protéines réceptrices olfactives et chacun de ces gènes exprime un seul type de protéine réceptrice olfactive [202]. Cependant sur le millier de gènes olfactifs, environ 400 d'entre eux sont actifs, les autres sont des pseudogènes. Les 400 récepteurs actifs ne sont pas les mêmes d'un individu à l'autre : en moyenne, deux individus diffèrent d'au moins 30% par des allèles de leurs récepteurs olfactifs [203].

Ainsi, le même composé n'active pas nécessairement le même ensemble de récepteurs chez des individus différents. L'absence d'un récepteur provoque une perception tout à fait différente et peut parfois même empêcher la détection de l'odeur : on parle dans ce cas d'anosmie spécifique [204,205]. Un exemple connu d'anosmie spécifique est l'odeur de verrat : l'odeur forte et désagréable que peut dégager la viande de porcs mâles non castrés au moment de la cuisson. Cette odeur n'incommoder pas tout le monde : selon les pays, 20 à 40% de la population ne la détecte pas alors que les autres la trouvent rédhibitoire [206]. Cette odeur provient essentiellement de deux molécules : le scatol (3-méthyl-indole) et l'androsténone (5-androst-16-ène-3-one). En raison d'un polymorphisme génétique, une grande proportion de la population ne perçoit pas d'odeur lorsqu'elle flaire la molécule d'androsténone, alors que d'autres perçoivent cette molécule comme désagréable (« sueur, urine ») et certains la perçoivent comme agréable (« douce, florale ») [207,208]. Un récepteur olfactif, nommé OR7D4, donne la capacité de détecter l'androsténone. Il existe plusieurs polymorphismes du gène codant ce récepteur. Certaines personnes sont porteuses d'un variant OR7D4 WM non fonctionnel, les rendant moins sensibles à l'androsténone [209,210].

Nous avons ainsi chacun une combinaison de récepteurs olfactifs qui nous est propre et par conséquent nous possédons une perception olfactive unique. Cette relation entre génétique et perception explique en partie la grande variabilité interindividuelle observée en matière d'odeurs, mais ce n'est pas le seul facteur de diversité de la perception olfactive.

2. Hormones et sexe biologique

Les femmes sont d'ordinaire meilleures que les hommes aux tests de détection, d'identification, de discrimination et de mémorisation des odeurs [211]. Cette meilleure sensibilité serait liée aux hormones mais également aux rôles sociaux, généralement

attribués aux femmes, les confrontant à un environnement olfactif plus riche [212] et les amenant à porter plus d'attention aux odeurs.

La sensibilité des femmes fluctue au cours de leur cycle menstruel [213,214], on observe un pic de sensibilité au milieu du cycle c'est-à-dire au moment de l'ovulation. Les femmes seraient ainsi plus réceptives aux signaux sociaux des hommes au moment où la probabilité de concevoir est plus élevée.

Les femmes enceintes présentent une réactivité émotionnelle plus forte face aux odeurs. Elles ressentent plus de dégoût pour certains aliments et certaines odeurs durant le premier semestre de la gestation [215,216]. Ce moment coïncide avec la phase d'immunodépression chez la mère dans laquelle le fœtus est le plus vulnérable. La réactivité accrue de la mère permettrait de protéger le fœtus de substances potentiellement nocives pour lui.

Par ailleurs, durant la ménopause, les femmes éprouvent parfois de l'anhédonie olfactive c'est-à-dire une baisse du plaisir olfactif [217].

3. Apprentissage

Nous apprenons à nommer et discriminer les odeurs durant nos rencontres quotidiennes avec celles-ci : à partir d'objets spécifiques comme les aliments utilisés lors de pratiques culinaires ou de l'utilisation de produits d'hygiène et d'entretien domestiques, mais également à partir de moments émotionnels (odeur de notre entourage ou des lieux fréquentés).

Cet apprentissage peut également se transmettre de manière pédagogique, en indiquant explicitement le nom ou les caractéristiques de l'odeur à l'enfant. Cela est peu pratiqué en Occident, mais il existe des cultures où les enfants ont une éducation olfactive, et particulièrement dans les populations vivant en environnements naturels où les modes de subsistance traditionnels sont centrés sur la chasse, la pêche, la cueillette et l'horticulture [218]. Les adultes apprennent ainsi aux enfants à trouver et identifier les plantes alimentaires et médicinales.

Dans notre société moderne, un adulte qui souhaite effectuer une carrière en lien avec l'olfaction, comme un parfumeur ou un œnologue, suit généralement une formation pour apprendre à reconnaître, identifier, nommer et décrire les odeurs. Cette formation consiste en un entraînement olfactif intense et régulier où les odeurs sont nommées et rattachées à un certain nombre de descripteurs olfactifs. Les descripteurs utilisés sont dépendants de l'organisme de formation. Les personnes formées, appelées communément « nez », ne sont pas plus sensibles que les non experts mais discriminent, mémorisent et décrivent mieux les odeurs grâce à leur apprentissage [219]. Elles savent également faire la distinction entre les différentes notes qui composent un parfum. Ainsi un expert ne donnera plus une évaluation subjective et émotionnelle comme décrit dans la partie « C.1.a.Langage courant » mais une description de l'odeur plus analytique du point de vue qualitatif.

4. Environnement olfactif et culture

Selon l'environnement olfactif dans lequel nous vivons, nous avons une expérience chimio-sensorielle différente qui impacte la façon dont nous percevons les odeurs. La fréquence d'exposition à une odeur ainsi que le contexte dans lequel elle apparaît vont avoir un impact sur l'appréciation de celle-ci, particulièrement durant la petite enfance. Ainsi notre histoire personnelle exerce une influence forte sur la manière dont nous percevons les odeurs en grandissant selon que ces odeurs sont rattachées à des souvenirs positifs ou négatifs.

Ainsi on observe de très fortes différences entre des populations ayant vécu dans des environnements olfactifs distincts. Les préférences alimentaires, par exemple, ne sont pas les mêmes suivant les pays : la consommation de fromage odorant très appréciée en Europe en particulier en France, ne l'est pas forcément dans le reste du monde. De même, pour des fruits comme le durian très apprécié en Asie du sud et peu apprécié dans d'autres régions du monde [220].

Il semble toutefois y avoir davantage de consensus concernant les odeurs désagréables que les odeurs agréables [221]. Cela s'explique par le fait que les sources dangereuses sont communes aux différentes populations (ex. le feu) alors que les éléments à approcher comme les aliments consommables diffèrent selon l'environnement.

Ainsi des personnes provenant de différentes cultures auront une perception olfactive très différente de par un héritage olfactif, des expériences et un environnement olfactifs divergents [218]. Des conditions environnementales différentes peuvent également avoir un impact.

5. Conditions environnementales physique et chimiques

Les conditions environnementales comme l'humidité, la température ou la pression barométrique modulent l'odorat. En effet, la température peut modifier les concentrations de composés gazeux et affecter les processus biologiques de l'odorat [222]. De plus, nous détectons mieux les odeurs dans des conditions humides par rapport à des conditions relativement sèches [223]. Par ailleurs, notre sensibilité olfactive est altérée dans un milieu à faible pression atmosphérique. Ainsi selon la saison ou le climat du pays dans lequel on se trouve nos performances olfactives ne seront pas forcément les mêmes.

L'air ambiant peut également contenir des produits agressifs pour notre olfaction. En effet, l'exposition aux polluants atmosphériques peut être associée à un dysfonctionnement olfactif [224]. Les polluants inspirés peuvent provoquer des dommages à chaque aire du trajet olfactif et créer des perturbations de la morphologie des tissus, des inflammations locales, des réponses au stress cellulaire, des lésions de la substance blanche et des pathologies vasculaires dans le bulbe olfactif, les zones corticales et sous-corticales frontales [225].

Il existe d'autres produits chimiques agressifs comme le tabac.

6. *Tabagisme*

Le tabagisme altère nos capacités olfactives. En effet, les fumeurs sont six fois plus susceptibles de présenter un déficit olfactif que les non-fumeurs et cela dépend de la dose de tabac consommé quotidiennement et de la durée de consommation [226]. Ces déficits olfactifs ne sont pas observés chez les anciens fumeurs : cela suggère que les capacités olfactives peuvent être pleinement récupérées à la suite de l'arrêt du tabac.

Plusieurs mécanismes pourraient expliquer l'effet du tabac sur l'olfaction. L'exposition à la fumée du tabac entraînerait une apoptose (mort cellulaire programmée) plus importante des neurones olfactifs [226,227]. Ceux-ci se régénérant régulièrement, l'arrêt du tabac permettrait une restauration complète de l'épithélium olfactif.

Un mécanisme supplémentaire serait la métaplasie squameuse, c'est-à-dire un remplacement de l'épithélium des voies respiratoires, dont fait partie l'épithélium olfactif, par un épithélium plus épais appelé épithélium squameux. Cet épaississement de la muqueuse olfactive permet de protéger notre organisme des agressions du tabac, par conséquent, elle rendrait l'accès aux récepteurs plus difficile pour les molécules odorantes. La métaplasie squameuse diminue dans les 6 mois suivant l'arrêt du tabac.

D'autres phénomènes pourraient potentiellement expliquer la baisse des capacités olfactives chez les fumeurs comme une inflammation naso-sinusienne, une réduction de la taille des bulbes olfactifs et une diminution de la quantité de matière grise dans le gyrus olfactif (tissu associé au cortex olfactif primaire) [228].

7. *Vieillesse*

À partir de 50 ans, on observe un déclin progressif des capacités olfactives (sensibilité, discrimination et identification) appelé presbyosmie. La presbyosmie concerne la moitié des personnes âgées de 65 à 80 ans et trois quarts des personnes de plus de 80 ans [229]. Ce déclin cognitif global est provoqué, entre autres, par la dégradation de la muqueuse olfactive, une ossification de la lame criblée, l'endommagement des neurones dans le bulbe olfactif et à la baisse d'activité de certaines structures cérébrales traitant les odeurs [230].

Dans certain cas, le vieillissement olfactif peut également être dû à une maladie chronique ou à des médicaments

8. *Maladies*

Différentes maladies peuvent avoir un impact négatif sur nos capacités olfactives : certaines maladies infectieuses comme la Covid-19 [231,232], des maladies neurodégénératives telles que la maladie d'Alzheimer [233] et la maladie de Parkinson [234] ou encore des maladies psychiatriques comme la schizophrénie [235] ou la dépression [236].

De plus, chez la population vieillissante, on observe une anhédonie olfactive [237] : les odeurs agréables le sont de moins en moins alors que le jugement hédonique des odeurs neutres et négatives ne change pas. Cela pourrait s'expliquer par l'importance de continuer à détecter les dangers potentiels.

Enfin, dans certaines formes de cancer, des médicaments peuvent également altérer notre odorat, notamment les agents chimiothérapeutiques [238–240].

9. Émotions

Nous avons vu dans la partie précédente que les odeurs sont de puissants déclencheurs d'émotions, mais les émotions peuvent également impacter notre perception de l'odeur.

En effet, l'état émotionnel d'un individu module ses capacités olfactives [241]. Dans un contexte émotionnel, une odeur habituellement considérée comme neutre peut être perçue comme ayant la même valence que celle du contexte.

De plus, les sentiments de colère réduisent nos capacités d'identification des odeurs, contrairement au stress et à la peur qui améliorent notre capacité à détecter et identifier les odeurs [242] et cela dans le but de recueillir des informations pertinentes pour la survie de l'individu.

10. Familiarité

L'exposition répétée à une odeur peut changer la perception de celle-ci [216], par exemple une odeur désagréable peut ainsi devenir de moins en moins désagréable avec le temps si elle sentie régulièrement [243].

Ce phénomène est important à prendre en compte lors de la conception d'un protocole expérimental. Une description émotionnelle donnée en début d'expérience sur un odorant n'est peut-être plus pertinente en fin d'expérience si le ou la participante a été exposé·e de nombreuses fois à cet odorant.

11. Appétit

Notre appétit influence la valeur émotionnelle que l'on attribue à une odeur de nourriture. De fait, lorsque l'on a faim les odeurs alimentaires nous semblent alléchantes et nous font saliver alors que lorsqu'on est rassasié ces odeurs peuvent nous dégoûter, voire nous donner la nausée. On nomme « alliesthésie alimentaire », le fait qu'une même odeur change de valeur hédonique en fonction de notre état de satiété [244].

Il existe un autre facteur majeur dans l'appréciation des odeurs alimentaires appelé « satiété sensorielle spécifique ». Elle concerne la diminution de la valence émotionnelle des odeurs des aliments consommés il y a peu de temps ; elle nous permet de cette façon de varier notre alimentation [180,245]. Chez certaines personnes, comme les personnes âgées, cette satiété sensorielle spécifique est diminuée et peut entraîner une malnutrition due au manque de variété du régime alimentaire.

Il est également intéressant de savoir que la faim améliore notre détection d'odorants, que l'odeur soit alimentaire ou non [246].

En résumé, nous avons vu dans cette section que la perception olfactive résulte de la fixation de molécules du milieu extérieur sur les récepteurs olfactifs situés dans la cavité nasale et d'un traitement cérébral complexe impliquant des aires sensorielles primaires et des régions émotionnelles et mnésiques. L'olfaction est un sens particulier, complexe, subjectif et sujet à une forte variabilité individuelle. Les perceptions qui en découlent sont dépendantes de multiples facteurs dont l'environnement dans lequel nous vivons, notre génome, notre vécu et notre sensibilité émotionnelle.

Dans ce cadre, le but principal de cette thèse est de mieux comprendre le lien entre les émotions et l'olfaction. Pour cela nous chercherons à découvrir comment s'expriment les émotions au niveau neurophysiologique avec des stimuli olfactifs. Pour atteindre cet objectif, il est primordial de prendre en compte la variabilité individuelle. Nous n'allons malheureusement pas pouvoir prendre en compte tous les facteurs de variations dans nos analyses mais nous pourrions en prendre le maximum en fonction des différents facteurs présents dans les données étudiées. Afin d'étudier au mieux la complexité de l'olfaction et des émotions nous proposons de répondre à cette problématique à l'aide d'outils provenant du domaine de l'intelligence artificielle.

III. L'intelligence artificielle comme outil d'analyse de la relation émotions et odeurs

Les résultats des études passées sur l'expression des émotions en matière d'odeurs utilisent des tests statistiques identifiant des tendances [247–249] (p. ex : le rythme cardiaque augmente avec la joie et diminue avec la relaxation), des tests de corrélations [150,250,251] (p. ex : l'activité de l'amygdale est corrélée avec l'éveil émotionnel) ou encore de classification donnant un pourcentage de précision des prédictions [252–256] (p. ex : un réseau de neurones arrive à prédire à 72% la valence d'une molécule à partir de sa chimie).

À ce jour, peu d'études ne se sont intéressées à la variabilité individuelle, pourtant très présente en matière d'odeurs et d'émotions. Les tests statistiques utilisés considèrent des paramètres comme les moyennes pour valider ou non une hypothèse commune à toute la population, ou comparent des populations préalablement déterminées. Les analyses de classification cherchent une règle générale qui sépare les différentes classes et qui fonctionne sur l'ensemble de la population étudiée. Cependant, nous avons vu que de nombreux facteurs influencent nos perceptions émotionnelle et olfactive : moyenniser des mesures aussi hétérogènes ou chercher un modèle général qui discrimine des perceptions olfactives et émotionnelles différentes, signifie ignorer cette variabilité individuelle, ce qui représente une véritable limite en termes d'avancées scientifiques dans le champ des odeurs et des émotions.

La présente thèse de doctorat va tenter d'incrémenter les paradigmes existants en intégrant, comme moyens d'analyse complémentaires, des méthodes issues de l'intelligence artificielle qui permettent de générer des modèles descriptifs intelligibles et qui prennent en compte et décrivent cette variabilité interindividuelle.

A. Vocabulaire et définitions

Les termes intelligence artificielle, data science, data mining, machine learning, classification et clustering sont difficiles à définir car ces disciplines sont communes à de nombreux domaines scientifiques, incluant les statistiques ou les bases de données, qui peuvent avoir un regard différent sur leur nature et leur utilisation. Ces termes ne seront pas utilisés de manière interchangeable dans ce manuscrit. Nous allons ainsi préciser la façon dont nous les définissons dans le cadre de cette thèse.

L'**intelligence artificielle** (IA) peut être définie comme un ensemble de capacités, que peut avoir une machine, comme : la résolution d'un problème, l'extraction et la représentation d'informations, l'apprentissage, la décision et l'amélioration elle-même [257]. C'est donc une discipline très vaste qui englobe, entre autres, la robotique, les systèmes multi-agents, les environnements intelligents, et la data science.

Nous nous intéresserons ici à un champ plus restreint de l'IA appelé **Data Science**. La data science (ou science des données) est un terme générique utilisé pour décrire les différents aspects du traitement des données (collecte, nettoyage, analyse et visualisation des

données). Elle est utilisée dans le but d'extraire des informations significatives et exploitables.

Le machine learning et le data mining font partie de la data science, ils se différencient par des objectifs distincts même si les frontières entre ces disciplines ne sont pas toujours aussi évidentes que présentées ci-dessous.

L'**apprentissage automatique supervisé** (ou supervised machine learning) s'applique sur un jeu de données étiqueté et consiste à créer un modèle de prédiction global à partir d'exemples annotés. Lorsque la variable à prédire est qualitative, on réalise une classification et s'il s'agit d'une variable quantitative (ensemble continu de réels), on parle d'une régression.

La **classification** s'utilise sur un ensemble des données découpé en deux ou plusieurs sous-ensembles de données nommés classes. Elle consiste à créer un modèle qui prédit au mieux les classes des différentes entités présentes dans les données.

L'**apprentissage automatique non supervisé** (ou unsupervised machine learning) apprend des modèles globaux à partir de données non étiquetées. Il peut être utilisé pour faire de la réduction de dimensions, de l'estimation de densité de distribution ou du clustering.

Le **clustering** (ou partitionnement de données) est une technique d'apprentissage supervisé qui permet de répartir l'ensemble des données en un certain nombre de groupes, nommés clusters, de manière à ce que les éléments compris dans un même cluster soient plus similaires entre eux qu'à ceux des autres clusters [258].

Le **data mining** (ou fouille de données) sera présenté ici comme domaine connexe mais distinct du machine learning. Il a ainsi pour but d'extraire à partir de données des nouvelles connaissances non triviales, implicites, potentiellement utiles et exploitables [259]. Il explore des modèles locaux en cherchant des relations entre les variables.

Le volume de données augmente continuellement en neurosciences du fait des questionnaires en ligne ou expériences faites à la maison et parce que de nombreuses expériences similaires déjà menées dans la communauté scientifique pourraient être étudiées ensemble sous forme de méta-analyse. Nous avons donc besoin d'un système capable de générer automatiquement une liste de connaissances extraites des données pour une meilleure compréhension de ces dernières. Les approches de data mining se révèlent particulièrement intéressantes pour notre sujet de recherche car elles permettent de développer des modèles descriptifs intelligibles à partir des données et peuvent prendre en considération les variations entre individus pour expliquer les données.

B. Processus de découverte de connaissances dans les bases de données

Comme expliqué précédemment, le data mining est un champ de la data science, ainsi son processus regroupe de nombreuses étapes allant de la collecte de données jusqu'à la visualisation des résultats.

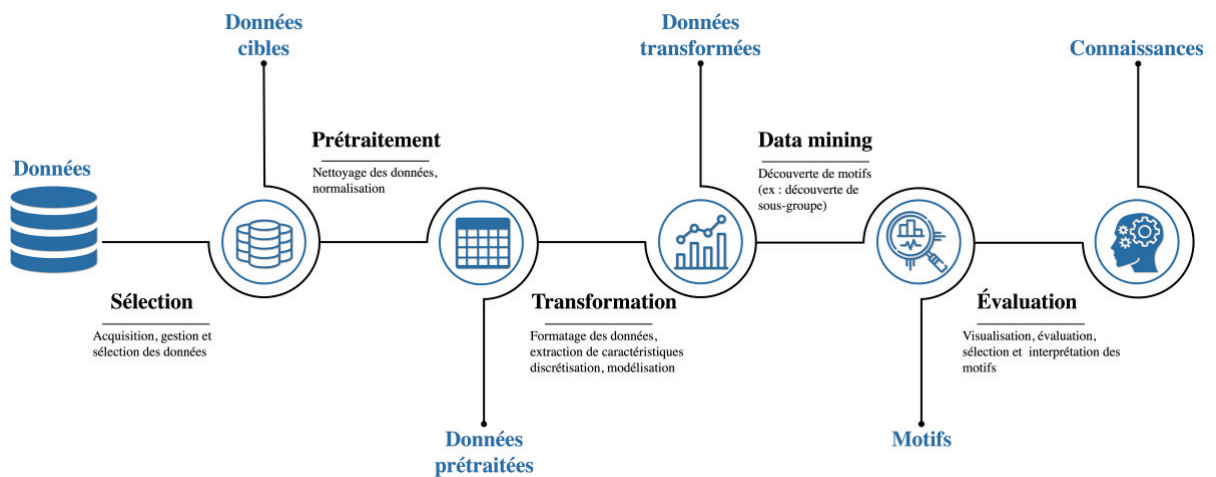


Figure 11 - Le processus « Knowledge Discovery in Databases » (KDD) décrit les 5 étapes de l'extraction de connaissances à partir d'une base de données.

Le pipeline utilisé classiquement en data mining comprend les étapes suivantes [260] :

- **Sélection** : Cette première étape sert à définir son objectif et à sélectionner le jeu de données correspondant. Ainsi elle peut comprendre l'acquisition de nouvelles données (ex : questionnaires, expérience en laboratoire), la sélection de données existantes, ou la combinaison de données provenant de différentes sources. Enfin, à partir des données collectées, elle permet l'extraction des données cibles intéressantes pour répondre à l'objectif fixé.
- **Prétraitement** : La seconde étape consiste à préparer et nettoyer les données sélectionnées. Cette tâche n'est pas à sous-estimer car c'est généralement la plus coûteuse. Le nettoyage porte sur la gestion des données redondantes, aberrantes, bruyantes ou manquantes. Il peut être nécessaire de supprimer les données non pertinentes, de corriger le bruit et d'estimer les entrées manquantes pour augmenter la qualité et la précision des résultats. La préparation des données peut également concerner la normalisation d'attributs de différentes échelles.
- **Transformation** : Les données utilisées sont rarement présentées sous un format adapté au traitement choisi. Ainsi cette troisième étape a pour but de transformer les données sous la forme requise par le ou les algorithmes de data mining utilisés dans l'étape suivante. Dans le cas d'un volume de données trop important pour les algorithmes de traitement choisis, l'extraction de caractéristiques permet de simplifier et transformer les données en caractéristiques significatives sur lesquelles le traitement pourra être réalisé. La sélection des caractéristiques les plus pertinentes nécessite une compréhension approfondie du domaine d'application [261]. Enfin, pour que les données soient converties au format attendu, il faut parfois discrétiser des descripteurs numériques pour les rendre nominaux, ou encore, modéliser des données sous forme d'un certain type de données comme des item sets, des séquences ou encore des graphes. Cette étape de transformation peut être considérée séparément de celle du prétraitement car elle dépend spécifiquement de l'analyse qui va suivre.

- **Data mining (ou exploration de données) :** Cette étape vise à concevoir ou utiliser des algorithmes efficaces identifiant des règles descriptives contenues dans les données transformées [262]. Les règles générées sont appelées motifs (ou patterns). Durant cette phase, il faut choisir le ou les algorithmes à utiliser mais également leurs paramètres de sorte à cibler aux mieux les motifs désirés.
- **Évaluation :** Cette dernière étape consiste à sélectionner et interpréter les motifs extraits précédemment avec la méthode de fouille de données. Durant cette étape, on réalise une synthèse, ainsi que des visualisations et des évaluations des motifs . On peut utiliser un algorithme pour valider la qualité des motifs et les motifs sont également évalués par l'utilisateur·rice qui juge leur pertinence par rapport à l'objectif choisi dans la première étape. Pour finir, les motifs gardés sont sujets à interprétations par les expert·e·s du domaine.

C. Type de données de base

Les types de données qui peuvent être intégrées à ces modèles sont multiples et variés. En guise d'exemples, nous allons présenter certains formats de données que nous avons utilisés dans ce travail de thèse. Plus précisément, nous allons prendre un exemple fictif de données et allons proposer trois modélisations possibles de ces données : matrice de donnée, graphes et itemsets.

Nous allons ainsi imaginer une étude cherchant à comprendre les préférences alimentaires des individus en fonction de leur nationalité. Un questionnaire a été envoyé en plusieurs langues dans de nombreux pays. Dans celui-ci, une liste d'aliments est présentée et pour chaque aliment, le participant indique s'il aime ou non le consommer régulièrement. Ainsi, les données brutes sont présentées Figure 12. Chaque ligne correspond à un individu et pour chaque individu, nous avons son identifiant, la liste des produits qu'il consomme ainsi que sa nationalité.



















<i>ID Participant</i>	<i>Aliments</i>						<i>Pays</i>
1							
2							
3							
4							
...

Figure 12 – Entête d'un jeu de données fictif. Chaque ligne correspond aux informations saisies dans le questionnaire par un participant. La première colonne indique l'identifiant unique du participant, les aliments présents sur une ligne correspondent aux produits qu'il consomme et la dernière colonne indique sa nationalité.

1. Données multidimensionnelles

La représentation la plus simple et la plus commune est la matrice de données, appelée également données multidimensionnelles [259,263]. Les données multidimensionnelles sont représentées sous forme d'un tableau de taille $r \times c$, avec r lignes et c colonnes. Les lignes correspondent en un ensemble d'entités (ou instances) indépendantes. Chaque entité est décrite par un ensemble d'attributs, appelés également champs, dimensions, variables ou caractéristiques. Chaque colonne correspond à un attribut. Ainsi chaque case donne la valeur d'un attribut pour une entité donnée.

Reprenons l'exemple Figure 12, ces données brutes sont déjà sous forme de données multidimensionnelles avec des valeurs binaires pour chaque aliment. On peut également agréger les données en les regroupant par pays. De la sorte, on calcule le pourcentage de personnes appréciant le produit de chaque pays et on obtient la matrice de données présentée Tableau 1. À partir d'un tableau comme celui-ci nous pourrions chercher les aliments fortement et spécifiquement consommés dans certains pays.

Tableau 1 - Modélisation du jeu de données fictif en matrice de données

Pays	Café (%)	Chocolat (%)	Fromage (%)	Poisson (%)	Porc (%)	Vanille (%)
France	51,6	56	96,1	40,1	42,1	82,3
Suisse	95,7	89,1	42,7	12,4	43,5	96,1
Allemagne	63,2	70,3	32,7	36,4	72	75,3
Portugal	19,5	52	42,8	86,3	56,3	74,9
...

À savoir qu'un attribut peut être de deux types différents :

- 1) Il peut être quantitatif c'est-à-dire avoir des valeurs numériques (ou continues) comme c'est le cas pour le pourcentage de chaque aliment Tableau 1.
- 2) Il peut être qualitatif (ou catégoriel) c'est-à-dire qu'il peut prendre des valeurs discrètes non ordonnées (ex. colonne « Pays »).

Tous les algorithmes ne prennent pas en compte des données multidimensionnelles mixtes comme le Tableau 1. En effet, certains algorithmes de data science acceptent seulement des attributs quantitatifs ou qualitatifs. Il faut choisir un outil adapté à son jeu de données ou parfois discrétiser les valeurs d'un attribut qualitatif.

Par ailleurs, tous les jeux de données brutes ne sont pas directement présentés sous ce format-là mais une modélisation en données multidimensionnelles peut généralement se faire via l'extraction de caractéristiques. Ainsi un texte peut être transformé en un tableau présentant les occurrences de termes génériques.

Encadré 2. Notations

Théorie des ensembles

Symbole	Signification	Exemple
{}	Ensemble d'éléments non ordonnés	$A = \{1, a, item, 2.5\}$
	Cardinalité d'un ensemble	$ A = 4$ car A contient 4 objets
\	Différence ensembliste	$A \setminus \{item, 1\} = \{a, 2.5\}$
	Tel que	L'ensemble des entiers compris entre 5 et 50 exclus : $\{x \in \mathbb{N} \mid 5 < x < 50\}$

Logique propositionnelle

Une proposition est une formule qui possède une valeur de vérité, elle est soit vraie soit fausse.

La proposition A = « $1+1=2$ » est Vrai

La proposition B = « $1>2$ » est Fausse.

Symbole	Signification	Exemple
\rightarrow	Implication	$A \rightarrow B$ signifie que si la proposition A est vraie, B est vraie.
\neg	Négation logique	Si A est vraie alors $\neg A$ est Fausse et si A est fausse alors $\neg B$ est Vraie
\wedge	Conjonction (ET logique)	$A \wedge B$ est Vraie si et seulement si A et B sont vraies
\vee	Disjonction (OU logique)	$A \vee B$ est Vraie si au moins l'une des deux propositions A ou B est vraie

Exemple :

Ainsi $E' = \{(x, y) \in E \mid x \in V' \wedge y \in V'\}$ signifie que E' contient toutes les paires (x, y) comprises dans E qui vérifient la proposition suivante : x et y appartiennent tous les deux à V' .

2. Graphes

La modélisation en données multidimensionnelles suppose que les entités sont indépendantes les unes des autres. Cependant, les instances peuvent être liées les unes aux autres. Dans notre exemple, les pays frontaliers sont certainement plus proches entre eux en termes de préférences alimentaires par rapport aux autres pays.

Les graphes servent à modéliser des entités et les relations entre ces entités. Un **graphe** $G = (V, E)$ est composé d'un ensemble non vide et fini de nœuds V (aussi appelés sommets) et d'un ensemble d'arêtes E , chacune constituée d'une paire non ordonnée de nœuds (v_i, v_j) , qui représentent chacun une relation entre les nœuds v_i et v_j [259,263]. Lorsqu'une telle relation a lieu entre v_i et v_j , on dit que v_i et v_j sont voisins, ils peuvent être également qualifiés comme étant connectés ou adjacents. Un exemple de graphe avec 7 sommets et 11 arêtes est illustré Figure 13.a.

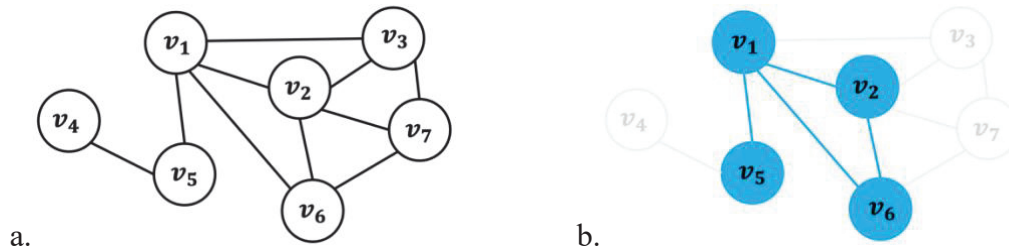


Figure 13 - a. Graphe connexe $G_1=(V,E)$, b. Sous-graphe du graphe G_1

Un graphe $G' = (V', E')$ est un **sous-graphe** de $G = (V, E)$ si $V' \subseteq V$ et $E' = \{(x, y) \in E \mid x \in V' \wedge y \in V'\}$ (voir encadré 2). Autrement dit, un sous-graphe G' de G est constitué d'un sous-ensemble des sommets de G et de toutes les arêtes présentes dans G entre les sommets V' de G' . Ainsi les sommets et arêtes bleus Figure 13.b constituent un sous-graphe de G_1 .

On nomme chemin ou chaîne, une suite d'arêtes permettant de relier deux nœuds. Dans la Figure 13.a, l'arête qui relie le nœud v_3 à v_1 , ainsi que celle reliant le v_1 à v_5 forment le chemin le plus court entre v_3 et v_5 . Un graphe $G = (V, E)$ est dit **connexe** s'il existe un chemin allant de x à y pour tous les sommet $(x, y) \in E$. Par conséquent, le graphe G_1 de la Figure 13 est connexe alors que le graphe G_2 de la Figure 14 ne l'est pas car il n'existe pas de chemin permettant de relier v_8 et v_2 par exemple.

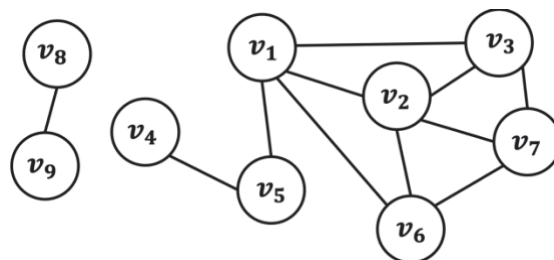


Figure 14 - Graphe non-connexe G_2

Certains graphes ont des spécificités en plus. Dans certains cas, un ensemble d'attributs peut être associé à un nœud, on parle dans ce cas de graphe à sommets attribués. Un **graphe à sommets attribués** $G = (V, E, A)$ est défini par un ensemble de sommets V , un ensemble d'arêtes entre les sommets $E \subseteq V \times V$, et un ensemble d'attributs A qui décrivent les sommets. La Figure 15 montre un exemple de graphe attribué : un nœud représente un pays, une arête relie deux pays limitrophes et les attributs indiquent les préférences alimentaires de chaque pays avec le pourcentage de consommation de chaque aliment.

À partir de cette modélisation, on pourrait voir si les habitants de pays proches spatialement ont des goûts similaires. En effet, il pourrait être intéressant de localiser des consommations similaires et spécifiques de certains aliments.

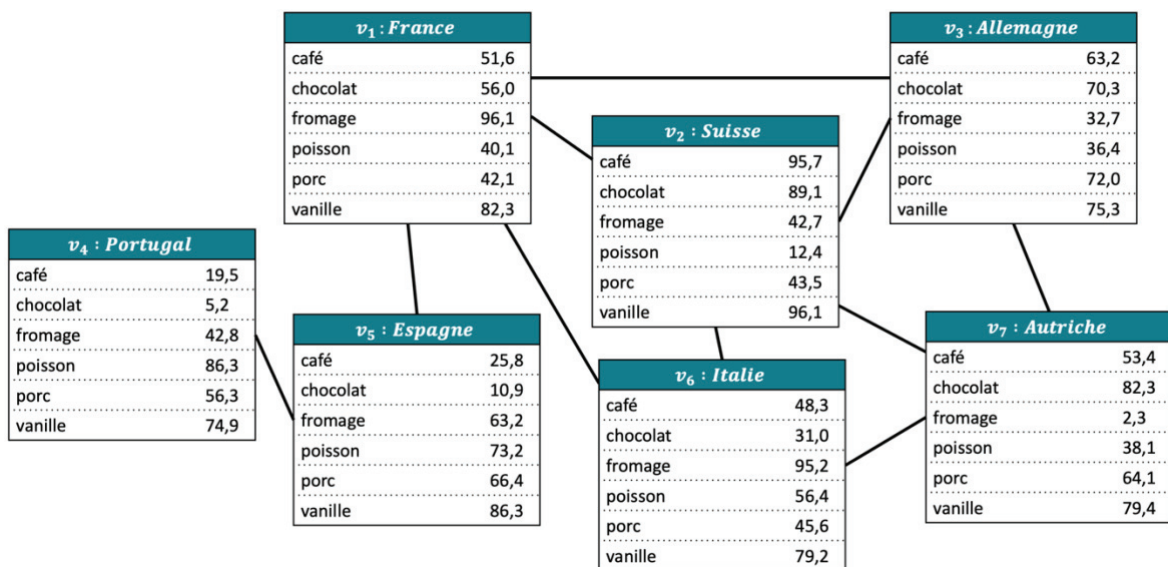


Figure 15 - Modélisation du jeu de données fictif en graphe attribué G_3 . $G_3 = (V, E, A)$ avec $A = \{\text{café, chocolat, fromage, poisson, porc, vanille}\}$

3. Itemsets

Les données binaires comme celles présentes dans la Figure 12 peuvent également être modélisées sous forme d'itemsets. Cette modélisation permet de s'intéresser à la fréquence à laquelle plusieurs objets coexistent.

Un **itemset** $I = \{i_1, i_2, \dots, i_n\}$ est un ensemble d'éléments appelés items. L'ensemble des éléments I peut désigner, par exemple, la collection de tous les produits consommés par une personne. Chaque ligne du jeu de données brut présentée en Figure 12, est appelée transaction.

Une **transaction** est composée d'un identifiant de transaction unique $t \in T$, et d'un itemset I qui lui est associé [263]. Dans notre exemple, l'identifiant t correspond à l'ID du participant. Pour l'individu $t=1$, on a l'itemset suivant : $\{\text{café, chocolat, fromage, vanille}\}$.

Les 4 premières transactions seraient les suivantes :

1	{café, chocolat, fromage, vanille}	France
2	{poisson, porc, vanille}	Espagne
3	{café, chocolat, vanille}	Suisse
4	{fromage, poisson, porc, vanille}	Espagne

Ainsi à partir de la liste de transactions ci-dessus, on peut, par exemple, chercher des ensembles de produits consommés spécifiquement et fortement dans un pays ou un ensemble de pays.

D. Méthodes utilisées

1. Découverte de sous-groupes

Dans le champ du data mining, nous nous intéresserons plus spécifiquement dans cette thèse à la découverte de sous-groupes et à l'exploration de modèles exceptionnels.

La découverte de sous-groupe (SD) [264–266] a pour objectif de trouver des liens interprétables entre différentes caractéristiques (variables descriptives) et une variable d'intérêt prédéfinie (variable cible). Ainsi, l'approche permet d'identifier et de décrire des régions dans l'espace des données d'entrée pour lesquelles la distribution observée d'une variable cible est significativement différente de celle du jeu de données dans son ensemble [267]. Les partitions de données descriptives trouvées sont nommées « sous-groupes exceptionnels ». Le caractère exceptionnel et l'intérêt d'un sous-groupe sont évalués par une mesure de qualité ; celle-ci compare le nombre d'occurrences de la ou des valeurs cibles dans le sous-groupe par rapport au reste des données. Il existe de nombreuses mesures de qualité différentes comme la Weighted Relative Accuracy (WRAcc) [268], le F-Score [269], les coefficients de Jaccard [270], l'informeness [271], le lift [272], etc.

Pour pouvoir découvrir les sous-groupes exceptionnels, l'algorithme utilisé doit explorer efficacement l'espace de recherche afin de donner des résultats qualitatifs dans un temps raisonnable. Les algorithmes **exhaustifs** explorent l'ensemble de l'espace de recherche et garantissent de trouver les meilleurs sous-groupes. Même si ces algorithmes exploitent de nombreuses techniques d'élagage, leur application s'avère limitée pour des jeux de données avec beaucoup de caractéristiques. Pour les données avec beaucoup de dimensions, les algorithmes **heuristiques** sont proposés, ceux-ci explorent le jeu de données de façon opportuniste et stratégique (ex : échantillonnage [273–275], algorithmes évolutionnaires [265,276,277]) afin de renvoyer les meilleures solutions trouvées en un laps de temps relativement court. Ils ne peuvent, par contre, pas garantir que celles-ci soient les meilleures de tout le jeu de données en réalité. Il existe également des algorithmes « anytime » [271,278] qui peuvent être interrompus à tout moment : ils produisent progressivement des sous-groupes dont la qualité s'améliore au fil du temps et si suffisamment de temps a été alloué, la découverte des meilleurs sous-groupes est garantie.

Même pour répondre à une problématique similaire, selon le type de jeux de données utilisés, ce n'est pas toujours le même algorithme qui sera utilisé.

Les paramètres d'entrée minimums requis pour un algorithme de découverte de sous-groupes sont généralement :

- 1) Les **variables descriptives** du jeu de donnée et leurs valeurs.
- 2) La **variable cible** et ses valeurs.
Lorsque la variable cible est discrète, il faut également préciser à partir de quelle classe de la variable cible l'on souhaite effectuer les sous-groupes.
Lorsque la variable est numérique, il faut parfois préciser si on s'intéresse à des intervalles, des valeurs élevées, des valeurs basses ou des égalités.
- 3) La **mesure de qualité** désirée si l'algorithme en propose plusieurs.
- 4) Un **seuil** permettant un filtrage des résultats. Celui-ci peut être une valeur de qualité minimale requise ou le nombre de sous-groupes souhaité. Dans le dernier cas l'algorithme renvoie les meilleurs résultats trouvés.

En sortie, l'algorithme renvoie des informations pour chaque sous-groupe trouvé contenant autre-autres :

- 1) Une **description caractéristique** du sous-groupe : elle précise les propriétés que doit satisfaire le sous-ensemble de données correspondant au sous-groupe.
Lorsqu'elle est écrite en langage propositionnel, elle consiste généralement en une conjonction de conditions sur les caractéristiques descriptives qui décrit précisément le sous-ensemble de données correspondant au sous-groupe.
- 2) La **qualité** du pattern via une ou plusieurs mesures de qualité, ainsi que d'autres indications sur le motif comme le support (nombre d'occurrence de la valeur cible).

2. EMM

L'exploration de modèles exceptionnels (EMM) [267] peut être considérée comme une extension de la découverte de sous-groupes. Alors que la découverte de sous-groupes ne concerne qu'une seule variable cible, l'EMM permet de gérer des cibles plurielles. L'EMM sert à identifier des partitions de données pour lesquelles un modèle ajusté sur les variables cibles est significativement différent de ce même modèle appliqué à l'ensemble du jeu de données.

Dans le cadre de cette thèse, nous avons estimé que les méthodes présentées précédemment seraient des candidates idéales pour étudier la variabilité émotionnelle et neurophysiologique dans la perception olfactive. En effet, ces méthodes ont déjà fait leurs preuves dans le domaine de l'olfaction : elles ont été utilisées dans le but de faire le lien structure chimique-odeur [279,280], description-odeur [281] ou encore perceptions-odeur [282].

De plus, ce qui rend les techniques d'exploration de données descriptives particulièrement attrayantes dans le contexte général des neurosciences, est le fait qu'elles impliquent de découvrir des hypothèses à partir de données. Le fait de générer des hypothèses via un algorithme pourrait permettre d'éviter l'expression de biais cognitifs et de

découvrir des hypothèses encore insoupçonnées. Cela est intéressant dans le cadre de l'étude du système nerveux dont le fonctionnement n'est pas encore bien compris aujourd'hui.



MEDIATION SCIENTIFIQUE
Ma Thèse en 180 secondes

J'ai participé au concours MT180 durant lequel les doctorants doivent exposer et vulgariser leur sujet de recherche en 3 minutes. Le pitch dans lequel je présente une analyse IRM lors de la finale lyonnaise est consultable sur la chaîne youtube de l'Université de Lyon. Les données et résultats abordés dans la vidéo ne sont pas les mêmes que ceux présentés dans ce manuscrit. La méthodologie expliquée reste cependant la même.

Lien: <https://bit.ly/3A5XyQS>

Dans cette thèse, nous nous intéressons ainsi à la caractérisation des comportements exceptionnels dans des ensembles de données neurophysiologiques. Nous cherchons des sous-ensembles des données, chacun caractérisé par une description sur les attributs neurophysiologiques pour une perception donnée (émotion positive par exemple).

IV. Problématique et objectifs de la thèse

Les éléments présentés précédemment montrent l'importance de la composante émotionnelle dans la perception olfactive du fait de son rôle prépondérant dans l'alimentation, la communication sociale et l'alerte d'un danger. L'odorat est sujet à une grande variabilité inter-individuelle et plusieurs facteurs de variation peuvent expliquer cette diversité. Le problème de savoir comment ces affects olfactifs sont représentés dans le cerveau humain et comment la diversité interindividuelle peut expliquer ces réponses émotionnelles n'est encore que trop rarement abordé aujourd'hui. Pour le résoudre, nous sommes limités méthodologiquement, faute d'outils d'analyse adéquats. Les objectifs de la présente thèse de doctorat sont au cœur de cette problématique.

Dans un **premier objectif**, nous tenterons d'apporter des éléments de connaissance scientifique sur les **représentations des odeurs et de leur composante hédonique au sein du système nerveux**.

Dans un **second objectif**, nous intégrerons à l'approche neuroscientifique une approche méthodologique issue de l'intelligence artificielle et du data mining afin de fournir des **modèles descriptifs et explicatifs** des effets des odeurs agréables et désagréables sur le fonctionnement du système nerveux central et périphérique. Ces modèles considéreront la diversité inter-individuelle en matière d'odeurs et d'émotions. Les algorithmes sous-jacents permettent la découverte de sous-groupes et visent à rechercher des patterns d'activité qui capturent une caractéristique propre aux individus partageant la même expérience olfactive (ex. odeur agréable ou odeur stimulante).

Pour atteindre ces deux objectifs, nous avons structuré la thèse en 5 études.

La **première étude expérimentale** se focalise sur des mesures du système nerveux périphérique comprenant : le rythme cardiaque, la respiration, la température cutanée et la conductance cutanée. À partir de ces données, elle tente de décrire le caractère plaisant et l'intensité d'une odeur à travers les réactions physiologiques périphériques. Les différents états affectifs (ex. plaisir) ont été décrits par des patterns sous forme de conditions sur les attributs psychophysiologiques. Certains patterns générés sont cohérents avec les résultats antérieurs d'études en neurosciences, et d'autres permettent de formuler de nouvelles hypothèses sur la perception olfactive. Par ailleurs, la participation de chaque individu aux règles découvertes permet de comparer les patterns correspondant aux odeurs agréables, à ceux correspondant aux odeurs désagréables.

Une **deuxième étude expérimentale** s'attaque à des données d'imagerie cérébrale par résonance magnétique fonctionnelle (IRMf) à partir desquelles nous avons cherché à localiser les zones cérébrales impliquées dans le traitement du caractère plaisant et déplaisant de l'odeur. Pour cela, une analyse univariée classique a été complétée d'une analyse par data mining. Dans cette analyse, nous avons modélisé les données IRMf sous forme de graphes attribués et extrait des sous-graphes exceptionnels. Ceux-ci nous renvoient des patterns d'activité spécifiques des odeurs plaisantes et déplaisantes au sein de deux structures importantes dans le traitement des odeurs : l'amygdale et le cortex piriforme. Ici, la mesure

de la participation des individus à chaque pattern a été effectuée à l'aide d'une approche originale inspirée de la théorie des jeux.

Une **troisième étude expérimentale**, encore préliminaire, concerne une expérience d'électroencéphalographie (EEG). L'aspect temporel des réactions affectives est abordé par une analyse temps/fréquence des signaux EEG. Les données ont été modélisées sous forme de graphes attribués sur lesquels nous avons réalisé une extraction de sous-graphes exceptionnels. Cela nous a permis de décrire la valence et l'éveil émotionnel perçus à partir des enregistrements des électrodes frontales F3 et F4. Les patterns générés fournissent des descriptions en accord avec les études antérieures sur le sujet et donnent de potentielles explications sur les résultats de certaines études de classification des odeurs plaisantes et déplaisantes. La variabilité individuelle a également été quantifiée.

Dans **une étude méthodologique**, nous avons proposé aux chercheurs en sciences sensorielles un tutoriel d'extraction de connaissances à partir de données physiologiques périphériques. À partir de données issues de la première étude expérimentale de cette thèse, nous avons détaillé sous forme de protocole les éléments intéressants, et les précautions à prendre durant l'analyse. Le partage de ces observations et recommandations a été vulgarisé au mieux afin de pouvoir servir à une communauté novice en data science.

Dans une **dernière réalisation**, plus théorique, nous avons mené une réflexion sur la place de la data science en neurosciences. Celle-ci répond à un article de De Wijk et Noldus [283] à propos de la nécessité de mesurer l'émotion dans un cadre écologique, hors du laboratoire. Avec le développement des nouvelles technologies (montre connectée par exemple) il devient de plus en plus accessible d'aborder les défis futurs concernant l'analyse des émotions au cours d'expériences hors du laboratoire. Nous partageons ces réflexions concernant l'analyse de grandes quantités de données dans et en dehors du laboratoire et proposons l'utilisation de méthodes provenant de la data science pour relever ces défis.

Ces cinq études sont présentées sous forme de 5 articles scientifiques dans la prochaine section de ce manuscrit, et les hypothèses testées pour chaque étude expérimentale sont intégrées dans les publications. À noter que durant ma thèse, et notamment dans le cadre de la pandémie de Covid-19, j'ai participé à plusieurs études complémentaires. Ces articles, ainsi que le matériel supplémentaire des études principales, sont présentés en [annexe](#).

RÉSULTATS

ÉTUDE 1 : « Utilisation de la découverte de sous-groupes pour relier le caractère hédonique et l'intensité des odeurs aux réactions du système nerveux périphérique» (publiée)

Matériel supplémentaire : [Annexe 4](#)

Using subgroup discovery to relate odor pleasantness and intensity to peripheral nervous system reactions

Maëlle Moranges, Marc Plantevit and Moustafa Bensafi

Abstract—Activation of the autonomic nervous system is a primary characteristic of human hedonic responses to sensory stimuli. For smells, general tendencies of physiological reactions have been described using classical statistics. However, these physiological variations are generally not quantified precisely; each psychophysiological parameter has very often been studied separately and individual variability was not systematically considered. The current study presents an innovative approach based on data mining, whose goal is to extract knowledge from a dataset. This approach uses a subgroup discovery algorithm which allows extraction of rules that apply to as many olfactory stimuli and individuals as possible. These rules are described by intervals on a set of physiological attributes. Results allowed both quantifying how each physiological parameter relates to odor pleasantness and perceived intensity but also describing the participation of each individual to these rules. This approach can be applied to other fields of affective sciences characterized by complex and heterogeneous datasets.

Index Terms— Mining methods and algorithms, Pattern analysis, Physiological measures

1 INTRODUCTION

A prominent property of olfaction, the sense of smell, is its affective component [1]. This level of processing is primal to humans and determines if a stimulus is edible or poisonous, pleasant or not, arousing or not. Affective responses to smells can be expressed verbally (i.e. subjective component), and are always accompanied by peripheral nervous system reactivity (i.e. physiological component) [2], [3], [4], [5]. The main function of these physiological responses is to provide somatic signals to the brain in order to guide important behaviors, namely approaching or avoiding the odor source [6].

An important question in the field is how these affective responses to odors relate to physiological activity. Beyond the olfactory sphere, and in a general way, there are two main theories on the topic. The first considers simple affective dimensions characterized by approach and withdrawal behaviors [7], [8]. Affects in this case, are categorized along axes of pleasantness, arousal and dominance. It has been shown that each of these dimensions varies according to specific physiological channels [9]. This is the dimensional or biphasic theory of emotion. The other theory postulates a limited number of discrete basic emotions (e.g. sadness, fear, joy, surprise, anger, disgust) characterized by distinct physiological events [10]. Note that whatever the theory, these affective states or basic emotions involve neural networks belonging to the limbic system (in particular the amygdala), regions such as the insula, but also associative systems located in the prefrontal and orbitofrontal cortex [11].

If we come back to the sense of smell itself, some researchers postulate an organization of the affects induced by smells in terms of approach and withdrawal systems [12], while others consider that affects provoked by odors can be distinguished according to basic emotions such as sadness, fear, joy, surprise, anger or disgust [4] [13], [14],

[15]. The truth is probably at the crossroads of these two theories insofar as odors can induce pleasant hedonic responses, pleasure, and the basic emotion of joy, or unpleasant hedonic responses, displeasure and the basic emotion of disgust [3]. Whatever the point of view or theory, the vast majority of psychophysical studies in olfaction shows that the subjective experience of odors is often characterized by a pervasive pleasantness dimension (e.g. "this odor smells very good or is bad") as well as descriptions related to the intensity of the odor (e.g. "this odor is very strong"). (see [16][17]).

As mentioned above, one prominent issue in the field concerns the relation between such perceptual and affective attributes of odors and physiological responses: how does the pleasant tone of a smell, and its perceived intensity, relate to peripheral physiological activity? Although past studies have revealed some relationship between physiological responses and odor pleasantness, and odor intensity, the directionality and the nature of these relationships varied from one study to another, and within individuals as a function of the ongoing task. For example, whereas for some, the electrodermal response increases with odor intensity [18], for others it varies according to the unpleasantness [19] or olfactory quality [20]. Such relationships between odor pleasantness and physiology have also been investigated using heart rate variability [21] and results showed that it is possible to discriminate between pleasant and unpleasant odors. Furthermore, whereas passively smelling odors or judging their pleasant tone induces similar heart rate activity patterns, familiarity judgment performed on the same smells overwrites this autonomic response [22]. These differences between studies and tasks may be a reflection of both methodological differences across studies but also of the large intra- and inter-individual variability in odor pleasantness and odor

intensity [5], [23], [24]. This variability across people makes it difficult to examine the relationship between smell pleasantness and intensity, and physiology and it is therefore necessary and important to consider it. To do so, it becomes essential to imagine the relationship between subjective experience of smells (pleasantness, intensity), and physiology no longer as a bijective relationship with a single rule that will link the 2 entities, but as a relationship that can be diverse: here, each subjective experience can be associated with multiple forms of physiological patterns in different individuals depending on biological factors but also on the developmental trajectories of each and everyone.

It is important to note that, in general, the data collected in psychophysiological studies are characterized by their heterogeneity and complexity at different levels: i/ in terms of stimuli, past protocols can range from a few to several dozens of items, ii/ in terms of human participants, a great interindividual variability characterizes the subjective judgments reflected by a very large diversity in the evoked hedonic response for a large range of smells, iii/ in terms of physiological responses, multiple parameters and variables are measured in the same individual. In addition to this physiological variability, there is a variability in olfactory perception linked to factors such as age [25], gender [26] or the menstrual cycle [24]. However, to arrive at a thoroughgoing model of odor pleasantness and intensity and their physiological basis requires having some means of i/ assessing these physiological responses in their diversity and heterogeneity across people and in their complexity, and ii/ describing in an intelligible manner the rules linking these physiological parameters to the individual subjective odor-driven affective experience: in fact, manually generating multiple rules linking the subjective experience (in its diversity) and physiological responses is not manageable.

Standard statistical approaches enable tackling the problem of the relationship between the subjective affective experience of odors and their physiological underpinnings very often on a single parameter basis (e.g. by comparing mean responses for pleasant vs. unpleasant for each individual physiological parameter). Furthermore, whereas predicting modeling approaches (e.g. classification algorithm) enable considering multiple physiological parameters in a single analysis, their explanatory power remains rather low. The aim of the present study is therefore **to better understand the relationship between the subjective experience of odor pleasantness and of odor intensity, and their physiological foundations by using a computational exploratory approach that we intend to apply to the field of affective sciences.**

To achieve this aim, we propose to use a data mining approach based on a subgroup discovery (SD) algorithm. SD enables to find population subgroups that are statistically "most interesting" from a population of individuals (or items) and a property on individuals of interest. In that respect, we aim to obtain the largest possible subpopulations which present the most unusual distributional statistical characteristics, and which respect the property of interest. The dataset used in the current study was previously published [17] and involves 22 individuals who

smelled a total of 109 stimuli (varying in olfactory quality - *chemical, medicinal, floral, etc.* - and in chemical families - *alcohols, aldehydes, esters etc.*). Each odorant was rated along two main dimensions of interest, namely pleasantness and intensity. Finally, a total of 4 physiological channels were recorded: finger pulse frequency (FP), skin conductance (SC), skin surface temperature (ST) and abdominal respiration (AR). These physiological data are described by 7 numeric attributes: one attribute for FP, AR and ST and four attributes for the SC (latency, rise time, amplitude and number of events).

Although the present study did not have the choice of physiological parameters to use (the data come from a previous study), some methodological choices should be clarified. Indeed, considering the duration of the experiment, Licon et al [17] privileged physiological sensors easy to install (belt positioned on the clothes for the breathing and sensors on the hand for the other parameters). Although less sensitive than the electrocardiogram for measuring cardiac activity and in particular its variability, the finger pulse was chosen here because its placement on the volunteers' thumb is extremely simple. Finally, note that the current analysis further explored 3 other subjective dimensions that are rarely studied and for which we intend to examine their physiological underpinnings: odor-evoked relaxation, anxiety and stress.

Our population is therefore obtained from subjective descriptions (pleasantness and intensity) of 109 flasks by different individuals. The property of interest concerns the search for subgroups identified by conditions on physiological attributes (e.g. $FP > 1.0$) for which a subjective experience (e.g. unpleasant) is significantly more present than in the rest of the data set. Note that the discovery of the best subgroups is NP-Complete (see Theorem 4 in [27]). Therefore, tackling this problem with a correct and complete algorithm is not feasible in practice. Indeed, we started by using an exact algorithm named *FSSD* (Fast and efficient algorithm for Subgroup Set Discovery)[28]; in several days it does not converge towards the first best solution and therefore never calculates the following solutions. To circumvent this issue, we therefore used an SD algorithm called *monteclopi* [29] (Monte Carlo Tree Search type, MCTS). MCTS is a heuristic that provides a fast and not necessarily optimal response by randomly exploring the dataset and improving the most promising results found by building a tree accordingly. In this way, it gives us the best intervals found for a given target attribute (e.g. pleasant, neutral, unpleasant).

2 RELATED WORK

Up to now, the understanding of the physiological foundation of the subjective experience of odor pleasantness and odor intensity has been explored using classical statistics and supervised classification. According to our knowledge, exploratory subgroup discovery methods were very few, or not yet applied in this field of olfactory affective sciences. For classical standard statistical approaches, in the very large majority of studies, hedonic responses are compared across conditions (e.g. pleasant vs.

unpleasant) by considering statistics such as the average or the amplitude of the signal in a given time interval. Here, complexity and heterogeneity of the data (e.g., heterogeneity in perceptual and physiological responses) is not well considered. These statistical methods used include mainly non parametric tests (e.g. Wilcoxon, Kruskal Wallis) [30] and parametrical tests (Analysis of Variance, ANOVA [31], [32],[33], [34], [35], [36], [37] or even MANOVA [38], [39], [40], [41], [42]) depending on the study design, normality of data and/or sample size. The main message of these studies is that pleasant and unpleasant odors act differently on the activity of the peripheral nervous system, and that certain parameters such as skin conductance, respiration, heart rate and its variability, or skin temperature may be of interest to discriminate these hedonic classes [3], [18], [19], [20], [21], [43], [44]. or basic emotions evoked by odors [4], [13], [14], [15].

Machine learning methods have been successfully applied in the field of olfaction for specific research questions. For example, there has been a rapid rise in the use of machine learning methods for predicting the relationship between a molecule's structure (e.g., the physico-chemical space) and its odor (e.g., fruity, musky, the perceptual space) [45]. Following this line of research, the most recent deep learning methods have been applied to tackle this challenge. GoogleAI [46] proposed the use of graph neural networks for studying the structure-odor relationship. They showed that these graph neural models outperform methods on tasks such as random forest model and k-nearest neighbor model on tasks of relating odor perception to odorant chemistry. Despite these recent successes, some research questions remain challenging: studying the relationship between two spaces that are not stable as the chemical space, and are characterized by their variability is much more difficult. This is the case when the core of our research question is investigating the link between the physiological space and the hedonic space which are characterized by their diversity and could be qualified as non stable. In that context, data science approaches were used in past physiological studies using a large variety of machine learning algorithms such as artificial neural networks (ANN) [47], [48], [49], [50], [51], [52], support vector machine (SVM) [49], [50], [53], [54], [55], multivariate pattern classification analysis (MVPA)[56], k-nearest neighbor [57], decision trees [58], random forest [50] and linear discriminant analysis (LDA) [59], [60], [61]. These approaches enable predicting subjective affective responses from autonomic activity. However, in most cases, these algorithms are qualified as black box classifiers (e.g. SVM or ANN) since their interpretability power is rather low which is not ideal for affective scientists who seek to explain and understand the physiological underpinnings of affective responses. Note that contrary to standard statistical approaches, only few of these data sciences studies were applied to the field of olfaction [61].

In summary, whereas some classifiers such as decision trees or regressions can provide a better explanatory power, they are still not relevant to achieve our aim because these methods attempt to predict classes of items in the entire dataset whereas our objective is to determine

whether there are subsets of items in the entire dataset that are characterized by exceptional physiological properties. Thus, the present article proposes a data mining method (subgroup discovery) that will attempt to search for such knowledge in the data by extracting, in the whole dataset, subgroups of items (odor pleasantness and odor intensity) characterized by exceptional rules (patterns of specific physiological responses).

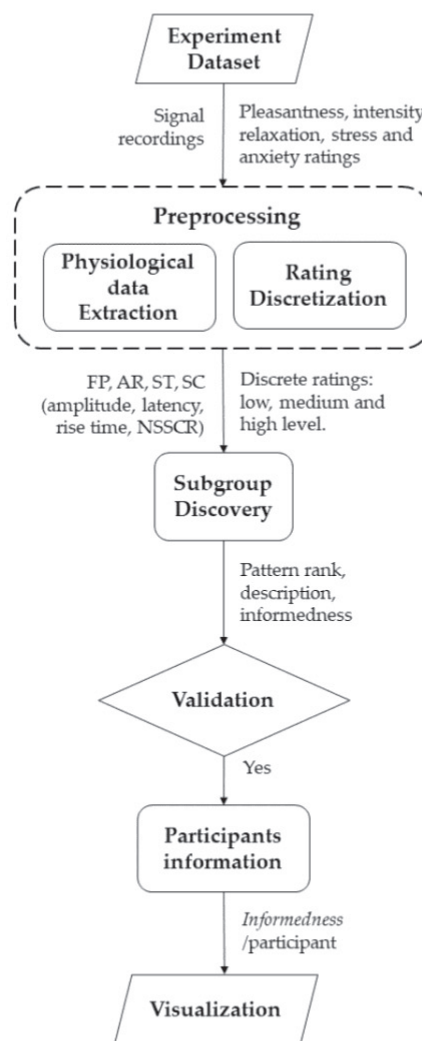


Fig. 1. Steps of the workflow. The raw data of the experiment are pre-processed (filtering, normalization, subtraction of the baseline) and the subjective ratings are discretized into 3 categories for each individual. Once the data is clean, the subgroup search describing the different perceptual dimension is performed. Then a statistical validation of the generated patterns is performed: the pattern is validated if its informedness is outside the confidence interval of the distribution of a bootstrapping of 1000 random draws. Finally, the extraction of information concerning the individual involvement of the subjects for each pattern is made and the visualization of all the results is accessible in different ways.

3 MATERIAL AND METHODS

In this section, we first present the psycho-physiological experiment from which the data was obtained. Secondly, we describe the preprocessing of physiological data, and that of perceptual data. Then, we present the process of

data analysis (from pattern generation using data mining algorithms, validation of these results by bootstrapping) to visualization of each individual in the observed patterns. An overview of the whole analytical process is shown in Fig. 1.

3.1 Experiment

The dataset comes from a previous study [17] which consisted in recording physiological responses to odorant molecules, together with psychophysical ratings. For a better understanding, we will summarize this experiment. Twenty-two subjects (11 men and 11 women) between the ages of 19 and 46 years (mean \pm SD, 32 \pm 10 years) participated in the study and each of them inhaled 109 different stimuli (105 odorants and 4 odorless trials containing only mineral oil). These odorant molecules were chosen so as to cover a large series of chemical families with different molecular weights and olfactory perceptual qualities. The odorants were presented in 15ml vials. These vials contained 5ml of odorants diluted in deodorized mineral oil and a scentless polypropylene fabric that absorbs stimuli to optimize evaporation and air/oil sharing.

During the experiments, participants were comfortably installed in a 7 \times 7 \times 4m room (room temperature of about 22 $^{\circ}$ C), in a semi-reclined seated position. This room was dedicated to the study of the sense of smell in that it was equipped with a ventilation system allowing a renewal of the ambient air. The Autonomic Nervous System (ANS) recording equipment included a total of four physiological parameters which were simultaneously and continuously recorded and displayed during the experiment: finger pulse (FP) frequency, skin conductance (SC), skin surface temperature (ST) and abdominal respiration (AR).

FP frequency was measured using a photoplethysmographic probe (3.2 cm/1.8 cm, LED photodetector) placed on the thumb of the non-dominant hand. Data were reduced to pulse rate, in beats per minute (bpm).

SC amplitude (in microsiemens: μ S) was recorded by two circular Ag/AgCl electrodes (diameter one cm) placed on the third phalanx of the forefinger and of the middle finger of the non-dominant hand. Data were reduced to latency (ms), rise time (ms), amplitude (μ S) and number of events.

ST was measured using a small ceramic-encapsulated metal-oxide semiconductor (9.5 mm length, two mm diameter). The thermistor, designed to operate from 0 $^{\circ}$ C to 50 $^{\circ}$ C, was placed directly on the first phalanx of the fourth finger. Data were reduced to skin temperature mean.

Changes in AR circumference were measured using a respiratory belt transducer (100 cm rest length, ten cm maximum elongation, 3.5 cm width), responding linearly to changes in length. Data were reduced to respiratory frequency, in cycles per minute (cpm).

Once ANS measures had stabilized, odorant trials were initiated. Each trial began with a white screen (for five seconds) followed by a written instruction, "Please prepare to smell", followed by a countdown, "3, 2, 1" (for five seconds).

Then, the investigator presented odorant stimuli about one cm below the participant's nose for three seconds.

Participants were instructed to sniff each flask for as long as it was presented, not to move and to focus their attention on the white screen in front of them during 30 seconds. Afterwards, they were asked to rate the odorant molecule along 5 dimensions: pleasantness, intensity, relaxation, stress and anxiety induced by the odorants, on a scale from 1 (not at all) to 9 (extremely). Participants could complete their ratings within a time window of 25 seconds. The 109 trials were divided in four sub-sessions or blocks (27 or 28 odorants/session), presented in random order. The duration of each block was around 30 minutes. Total session duration was around three hours.

It should be noted that in the original study by Licon et al [17], the same participants returned to the laboratory one day later for a second session. During this second session, participants were presented with the same series of 109 flasks, and their tasks were to rate, among others, each stimulus for the following 6 basic emotional states: joy, sadness, fear, anger, disgust, and surprise; neutral state and pleasure were also added. For more details, refer to the second session in Licon et al [17].

3.2 Physiological data and parameters

Physiological data were sampled and recorded at 256 Hz, then converted and amplified via an 8-channel PRO-COMP+ system (Thought Technology, Montreal, Canada) and displayed, stored, reduced and analyzed off-line. This device has already been used in several studies related to the senses and emotions in various fields such as psychiatry [62], food sciences [63] or neuroscience [6]. All physiological parameters were analyzed in the 30-s window after stimulus onset (using the ten seconds before odor presentation as baseline). Abdominal respiration (AR, cycle per minute or cpm), skin temperature (ST, mean temperature in $^{\circ}$ C) and finger pulse frequency (FP, in beats per minute or bpm) were normalized by subtracting baseline values (in the 10-sec period preceding stimulus onset) from the values in the 30-s window after stimulus onset. Note that for AR and FP, baseline values were multiplied by a factor

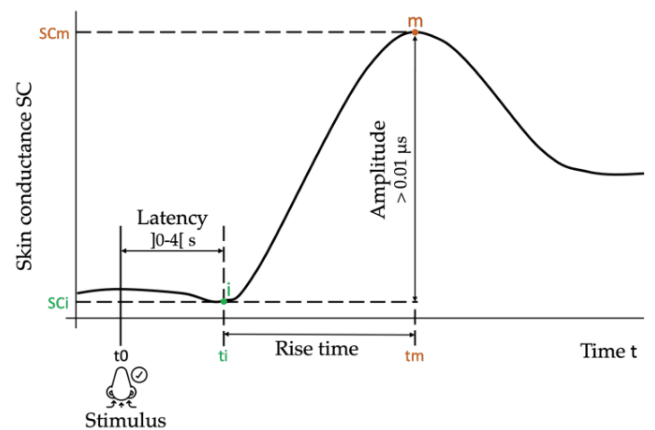


Fig. 2. Physiological attributes of the electrodermal response following stimulus presentation: latency (ms), rise time (ms), amplitude (μ S). t_0 corresponds to the time-point of odor presentation, i is the initiation point of the SC response and m is the maximum point of the SC response. The amplitude minimum is 0.01 μ S and the initiation point is sought within the first 4 seconds. The calculation of the 3 variables is as follows: Latency = t_0-t_i , Rise Time = t_m-t_i , Amplitude = SC_m-SC_i .

of 3 in order to be in the same time scale as that of the post-stimulus period (30 sec).

For Skin conductance (SC), in terms of signal analysis, SC has been smoothed by a Gaussian filter and analyzed through 4 variables. It should be noted that multiple SC responses can appear in the time between two trials. These multiple responses were considered by computing the number of skin conductance event (NSSCR) before and after the signal (as AR and FP), during baseline and after stimulus onset. We then focused on the first SC event following the olfactory stimulation by computing SC amplitude, latency and rise time, which are the most common SC parameters used in emotion and olfaction studies [19], [17], [49], [50], [64]. These SC parameters are depicted in Fig. 2. Note that the SC response begins in a time window of 4 seconds (maximum) after stimulus onset. In the case whereby no peak has been identified, the 3 attributes (latency, amplitude and rise time) are equal to 0.

3.3 Perceptual data and discretization

The ratings of each subjective scale were discretized using k-means [65] clustering ($k = 3$ classes; low, medium and high for pleasantness, intensity, relaxation, stress and relaxation). This discretization method has been chosen because it takes into account individual's valuation strategy which is not the case with more standard methods based on equi-depth and equi-width technics (see Fig. 3). This discretization has been applied to each participant independently. K-Means enable partitioning into k clusters so that the distance between intra-cluster points is minimized and the inter-cluster distance is maximized allowing partitioning the subjective data in three categories as different as possible. Note that if an individual did use less than 3 different ratings on all odorants for a given subjective dimension, then a discretization into 3 classes was not possible. In such cases, the individual was not included in the analysis for this specific subjective judgment (this case only happens for anxiety ratings of 2 individuals).

Fig. 3 illustrates an example of the different discretizations based on subjective ratings of a participant judging 6 different odorant molecules. A visualization of the discretization of each subject for each perceptual attribute is available in Fig. S1 (Supplementary file).

Note that to validate the classes generated by the k-means clustering, we ran the algorithm 100 times with different initializations for the same individual and a given subjective dimension. Results revealed a relative stability of the generated clusters (95.1% in average for all subjects for pleasantness, 95.9% for intensity, 95.8% for relaxation, 97.3% for stress and 96.3% for anxiety). We further validated these clusters using silhouette analyses in each individual and showed that for all dimensions, silhouette coefficient values were relatively high, which validate the consistency of clusters: pleasantness: [0.61-0.81], intensity: [0.64-0.88], relaxation: [0.66-0.93], stress: [0.67-1], anxiety: [0.63-1].

3.4 Subgroup discovery analysis

The above listed pre-processing ended in a dataset D analyzed through a subgroup discovery method described



Fig. 3. Example of classical discretizations and the k-means discretization ($k=3$) in 3 discrete variables: low, medium and high based on subjective ratings of a participant evaluating 6 odorants (ratings: 2,5,6,7,9,9). In literature, self-reported ratings are usually discretized by dividing the ratings in k groups with $k=2$ or 3. The methods used are generally equi-depth (division into k equal size intervals) [58], [66], [67] or equi-width (quantile division: k groups with the same number of values) [3].

below. The dataset D consists of the set O of m odorant molecules, the set S of n individuals, a set P of physiological attributes and a set E of subjective evaluations.

By convention, we denote by x_i , the i th attribute of the set of attributes X :

$$\begin{aligned} O &= \{o_1, \dots, o_m\} \\ S &= \{s_1, \dots, s_n\} \\ P &= \{p_1, p_2, p_3, p_4, p_5, p_6, p_7\} \\ &= \{FP, AR, ST, SC \text{ amplitude}, SC \text{ latency}, SC \text{ rise time}, \\ &SC \text{ NSSCR}\} \\ E &= \{e_1, e_2, e_3, e_4, e_5\} \\ &= \{Pleasantness, Intensity, Relaxation, Stress, Anxiety\} \end{aligned}$$

A pair of a given odorant and a given individual uniquely identifies a tuple in dataset D :

$$D: (o, s) \rightarrow P \times E$$

The value of the attribute $p \in P$ is denoted by $p(o, s)$. Similarly, the value (aka class) of the attribute $e \in E$ is denoted by $e(o, s)$. Note that for these attributes, there are 3 classes generated by the discretization process: low, medium, high. The set of tuples for an attribute e of class label c , are as follows:

$$D^{(e,c)} = \{(o, s) \in D \mid e(o, s) = c\}$$

An example of dataset is presented in Table 1a with only one perceptual attribute (pleasantness). In this example, $D^{(Pleasantness, medium)}$ corresponds to the "medium" class lines represented in bold in Table 1b. The other classes represented in italics are noted: $D^{(Pleasantness, \neg medium)}$. So, in terms of support: $|D^{(Pleasantness, medium)}| = 5$ and $|D^{(Pleasantness, \neg medium)}| = 4$.

We used a subgroup discovery approach to find a set of pairs, of individuals and odors that characterizes the specific class c against the other classes ($\neg c$) for a targeted perceptual attribute e . These subgroups are identified by an intent, i.e., a description. For this, we seek to identify the conditions on some physiological attributes strongly associated with a given subjective class (ex. Medium, Low, High). This can be formalized by a descriptive rule related to a subgroup.

TABLE 1
EXAMPLE OF DATASET

		Descriptive attributes							Target
a. Subject	Odor	FP	AR	ST	SC	SC	SC	SC	Class : Pleasantness
		(bpm ^a)	(cpm ^b)	(°C)	Latency (ms)	Rise_Time (ms)	Amplitude (μs)	NSSCR (nb events)	
s_1	o_1	0.34	-0.34	-0.04	2.98	3.02	0.2	1.0	Medium
s_1	o_2	-1.67	-0.34	-0.03	0.80	3.44	0.08	0.0	Low
s_1	o_3	2.0	0.0	0.020	0.0	0.0	0.0	1.0	Medium
s_2	o_1	-1.0	-1.0	-0.021	0.0	0.0	0.0	0.0	Medium
s_2	o_2	-0.33	-0.33	-0.05	0.0	0.0	0.0	0.0	Medium
s_2	o_3	-1.0	-0.33	-0.023	1.81	3.05	0.15	5.0	High
s_3	o_1	1.0	0.33	-0.05	1.81	3.36	0.11	0.0	High
s_3	o_2	3.67	0.67	0.06	2.46	1.81	0.03	4.0	Low
s_3	o_4	-6.33	0.33	-0.68	2.49	1.79	0.06	2.0	Medium

Example :
 e = Pleasantness
 c = Medium
 $r = SC_NSSCR \in [0,1]$

Legend :
 $cov(r)$
Bold $D^{(e,c)}$
Italics $D^{(e,-c)}$

b. Subject	Odor	SC NSSCR	Class : Pleasantness
s_1	o_1	1.0	Medium
s_1	o_2	0.0	<i>Low</i>
s_1	o_3	1.0	Medium
s_2	o_1	0.0	Medium
s_2	o_2	0.0	Medium
s_2	o_3	5.0	<i>High</i>
s_3	o_1	0.0	<i>High</i>
s_3	o_2	4.0	<i>Low</i>
s_3	o_4	2.0	Medium

c. Subject	Odor	SC NSSCR	Class : Pleasantness
s_1	o_1	<u>1.0</u>	Medium
s_1	o_2	<u>0.0</u>	Low
s_1	o_3	<u>1.0</u>	Medium
s_2	o_1	<u>0.0</u>	Medium
s_2	o_2	<u>0.0</u>	Medium
s_2	o_3	5.0	High
s_3	o_1	<u>0.0</u>	High
s_3	o_2	4.0	Low
s_3	o_4	2.0	Medium

d. Subject	Odor	SC NSSCR	Class : Pleasantness
s_1	o_1	<u>1.0</u>	Medium
s_1	o_2	<u>0.0</u>	<i>Low</i>
s_1	o_3	<u>1.0</u>	Medium
s_2	o_1	<u>0.0</u>	Medium
s_2	o_2	<u>0.0</u>	Medium
s_2	o_3	5.0	<i>High</i>
s_3	o_1	<u>0.0</u>	<i>High</i>
s_3	o_2	4.0	<i>Low</i>
s_3	o_4	2.0	Medium

a) Example of dataset for the subjective ratings of pleasantness (note that for FP, AR, ST and SC-NSSCR, each value is calculated by subtracting a baseline value from the physiological signal following odor presentation - see section 3.2; bpm means beats per minute and cpm means cycles per minute). b) Support for the “medium class”. c) Support for the pattern $r=SC_NSSCR \in [0,1]$. d) Explanation of informedness with the true positives in bold and highlighted and the false positives in italics and highlighted for the pattern r and the class c .

A descriptive rule r , denoted $r: d \rightarrow (e, c)$, is defined by a physiological description strongly associated to the class c of a targeted perceptual attribute e . The description is a set of intervals $r = \{[x_1, y_1], [x_2, y_2], \dots, [x_k, y_k]\}$, each being a restriction on the value of physiological attributes. The pairs (o, s) whose physiological attribute values belong to the intervals of description D are members of the coverage of D :

$$cov(r) = \{(o, s) \mid \forall i = 1 \dots k, x_i \leq p_i(o, s) \leq y_i\}$$

Let's consider the following pattern $r = "SC_NSSCR \in [0,1]"$ and the previous dataset this time illustrated in the Table 1c. $cov(r)$ corresponds to the gray lines, that means all the lines with the number of events is 0 or 1 (see underlined values).

We aim to discover subgroups where one target class for a given perceptual attribute is over-represented, the other entities being under-represented. For that, we are interested in the subgroups that maximize a quality measure. There are several quality measures [68], [69]. In this study, we focused on the *informedness* [70] which corresponds to the difference between the true positive rate and false positive rate of the pattern. This quality measure allows to identify subgroups that maximize coverage for one class target and minimize coverage for other classes. It is a trade-off between recall and precision.

$$informedness(r, e, c) = \frac{|cov(r) \wedge D^{(e,c)}|}{|D^{(e,c)}|} - \frac{|cov(r) \wedge D^{(e,-c)}|}{|D^{(e,-c)}|}$$

Let us take again the examples of the Table 1b and 1c: let us consider, for the target “medium”, the subjective evaluation of the “pleasantness”, and the pattern: $r = "SC_NSSCR \in [0,1]"$.

A new annotation of the dataset is depicted in Table 1d. Here, the lines in grey correspond to $cov(r)$, those in bold to $D^{(e,c)}$ and those in italics to $D^{(e,-c)}$. The lines in bold and highlighted illustrate the true positive, namely $cov(r) \wedge D^{(e,c)}$ and the bold and highlighted lines the false positives, namely $cov(r) \wedge D^{(e,-c)}$.

We therefore have the following quality score:

$$Informedness(r, pleasantness, medium) = 4/5 - 2/4 = 0.3$$

The objective of a subgroup discovery algorithm is therefore looking for physiological intervals which optimize *informedness*. Because the search space is huge, generating all the solutions is not possible, extremely time consuming and actually not feasible. We therefore opted for heuristic approaches which return the best solutions found for a given time budget. We use the *monteclopi* algorithm [29], which is based on Monte Carlo Tree Search (MCTS) [71] to explore the search space. Especially, *monteclopi* builds a search tree, which is expanded and updated at each iteration. Each path of the tree is described by a score determined by a quality measure. This algorithm is a combination of classic tree search implementations that uses exploration and exploitation. Exploration consists of randomly searching for new paths in order to find another solution. Exploitation allows the improvement of the most promising paths so as to maximize the measure of quality.

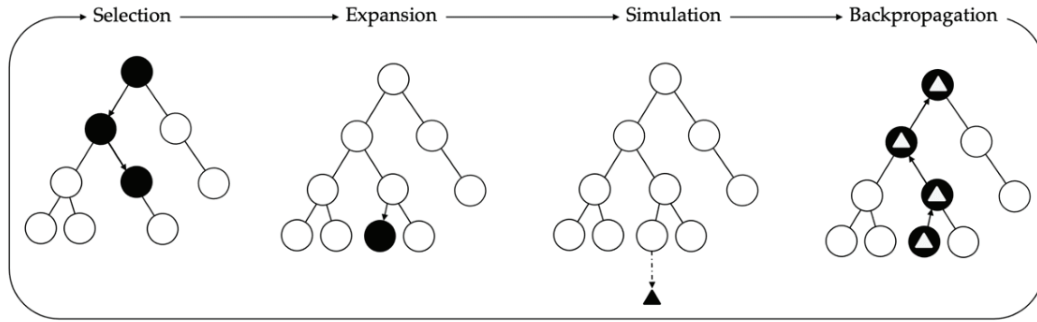


Fig. 4. Steps of the MCTS algorithm.

Monteclopi thus allows us to look for other good solutions while validating the quality of the current best. For that, at each iteration, the algorithm goes through four phases: selection, expansion, simulation (aka roll-out) and backpropagation (aka update). These four steps are illustrated in Fig. 4. Selection consists in selecting the unexplored child of a node from the tree already constructed. This node is developed in the expansion step and a new transition is added to the construct tree. The tree then simulates the value of this state: it successively performs random transitions from the extended node, until it reaches a terminal node. Then the backpropagation state allows to update the tree by adding the value of the new state and by correcting the value of each state upstream in the tree. These four steps are executed successively, until there is no more time budget or all the search space has been explored. At this point, the nodes with the best *informedness* are returned.

Monteclopi is given a time budget as a parameter (here 200 seconds) and it returns the 10 best patterns found at the end of the allotted time.

3.5 Validation of patterns

In order to validate the quality of a pattern, we performed bootstrapping by calculating the *informedness* of 10,000 groups for each pattern discovered, with the same support as the current pattern, drawn at random (with replacement between each draw). The aim is to validate the pattern if its *informedness* is outside the 95% of the random distribution. For this, we calculate the interval between the $100.\alpha$ and $100.(1-\alpha)$ percentiles of the bootstrap distribution with $\alpha = 0.025$. Only patterns with an *informedness* above the 97.5th percentile are kept, the others are rejected. An example of pattern validation is illustrated in Fig. 5. This interval-based validation makes it possible to avoid reporting subgroups indicating an *informedness* likely to be observed by a random subset of entities.

3.6 Participation of individuals in the generated patterns

Here, we calculated the number of individuals participating in a given pattern and we quantified how much they participate in it. To this end, we considered only the data of an individual when the *informedness* is positive. Let us consider s_i an individual, the data for this individual are noted:

$$D_{s_i} : (o, s_i) \rightarrow P \times E$$

For example, in Table 1, D_{s_1} corresponds to the first three rows in the table.

The *informedness* computed for a participant s for the rule r and the target (e, c) is:

$$\text{informedness}(r, e, c, s) = \frac{|\text{cov}(r) \wedge D_s^{(e,c)}|}{|D_s^{(e,c)}|} - \frac{|\text{cov}(r) \wedge D_s^{(e,\neg c)}|}{|D_s^{(e,\neg c)}|}$$

The higher the *informedness* for an individual, the more the individual participates.

The number of participants to the pattern r and the target c is:

$$Nb_{\text{participants}(r,e,c,s)} = |\{s \forall s \in S, \text{informedness}(r, e, c, s) > 0 \text{ is True}\}|$$

Let us consider the previous example and always the target “medium” for the “pleasantness” and the pattern: $r = \text{“SC_NSSCR} \in [0,1]\text{”}$. The number of participants for this rule and this target is computing from the *informedness* of each individual:

$$\text{Informedness}(r, \text{pleasantness}, \text{medium}, s_1) = 2/2 - 1/1 = 0$$

$$\text{Informedness}(r, \text{pleasantness}, \text{medium}, s_2) = 2/2 - 0/1 = 1$$

$$\text{Informedness}(r, \text{pleasantness}, \text{medium}, s_3) = 0/1 - 1/2 = -0.5$$

$$Nb_{\text{participants}(r,\text{pleasantness},\text{medium})} = \{s_2\} = 1$$

Thus, whereas Subject 1 and Subject 3 do not participate in the pattern and Subject 2 participates to the pattern at 100%.

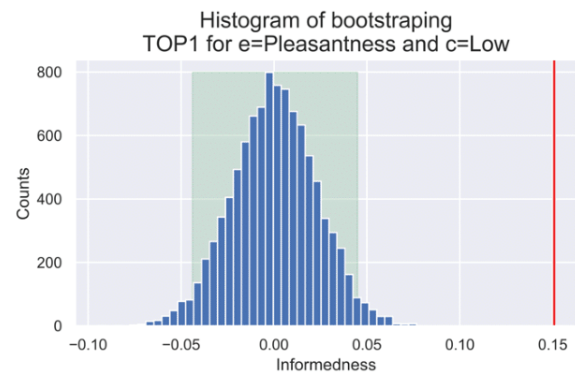


Fig. 5. Example of validation. In blue: the bootstrapping distribution. The rectangle around the distribution: the 95% percentile interval. The vertical bar: the informedness of the TOP1 pattern of the target “low” for pleasantness.

4 RESULTS

4.1 Overall description of results and quality

All 10 best descriptive rules of each perceptual attribute are presented in Tables S1-5 (Supplementary file). The set contains 150 patterns (top 10 descriptive rules, 3 classes, 5 perceptual attributes). The *informedness* of these patterns is between 0.042 and 0.151, which may be seen as not high but still outside the 95% confidence interval (for a better understanding of the bootstrapping procedure, see section 3.5 and especially Figure 5). As can be seen in Figure S2 (which shows the Top 1 validations of each category), for each pattern, the quality measure (vertical bar) is located beyond the upper limit of the confidence interval (rectangle around the distribution). These patterns are therefore significant.

The number of individuals participating in the patterns varies between 20% and 86.36%. Finally, whereas 83,33% of the rules contain only one interval ($\neq [-\infty; +\infty]$) (named as *Simple descriptive rules*; aggregated result shown in the Fig. 6, Fig. S8 and section 4.2), 16,66% of the rules contain multiple intervals (named as *Multidimensional descriptive rules*; presented in section 4.3).

For descriptive purposes, a specific terminology will be used to describe the different classes within each subjective dimension. For odor pleasantness, we will use the terms “unpleasant”, “neutral”, and “pleasant” for the respective classes “low”, “medium”, and “high”. For odor intensity, we will use the following notations: we will use the terms “not at all intense”, “midly intense”, and “very intense” for the respective classes “low”, “medium”, and “high”. Note that for the 3 other subjective ratings, the same notation as that for intensity will be used: “Not at all” + relaxing/stressful/anxiety-inducing for the “low” class,

“Mildly” + relaxing /stressful/anxiety-inducing for the “medium” class, “Very” + relaxing/stressful/anxiety-inducing for the “high” class.

4.2 Simple descriptive rules

Fig. 6 depicts the *simple descriptive rules*. These rules are represented in a combined visualization for odor pleasantness and odor intensity. For the same class (low, medium, high), one can have several rules whose interval concerns the same physiological attribute. For visualization purposes we presented the rule with the best *informedness*. All these selected rules are displayed on axes for each physiological attribute.

For pleasantness (Fig. 6a), it can be seen that pleasant odors are associated with a decrease in FP activity, and very small variations in skin temperature; combined with neutral odors, they decrease activity in most SC attributes. For neutral odors, they decrease SC activity (number of events) and skin temperature. For unpleasant odors, the activity is mainly focused on one parameter, namely SC which increases for 3 attributes (amplitude, latency and rise time). For intensity (Fig. 6b), not at all intense odors decrease AR, increase ST and decrease SC (number of events). Moreover, there is a gradual increase of SC activity from mildly intense to very intense for 3 attributes (amplitude, latency and rise time). Finally, very intense odors induced an increase in SC activity for the attribute number of events.

Mixed results were observed for relaxation, stress and anxiety (Fig. S7a-c). Not at all relaxing odors induced few variations in AR, a decrease in ST and an increase in SC latency and rise time. Mildly relaxing odors induced a decrease in FP and AR activities and small increases in SC amplitude and rise time. Very relaxing odors induced an

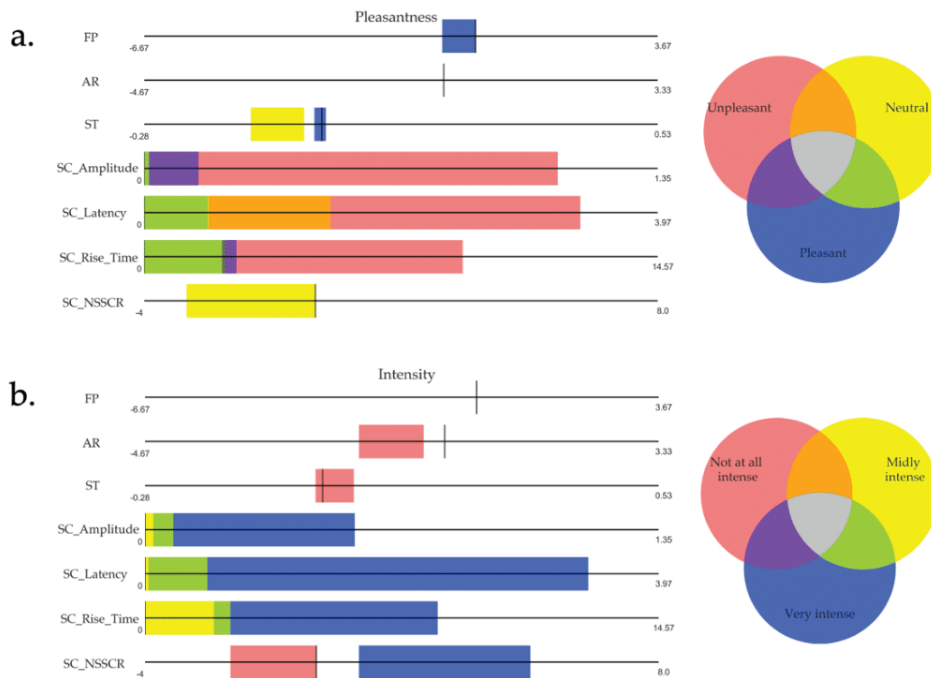


Fig. 6. Visualization of simple rules for odor pleasantness (a) and odor intensity (b). Each vertical line corresponds to the value zero (0).

TABLE 2
PERCENTAGE OF MULTIDIMENSIONAL RULES FOR EACH AFFECTIVE DIMENSION AND CLASS

Classes	Pleasantness (%)	Intensity (%)	Relaxation (%)	Stress (%)	Anxiety (%)
Low	0	0	0	0	0
Medium	50	60	0	60	70
High	20	0	0	0	0

The other rules are one-dimensional, they are called simple rules and are described in section 4.2.

increase in AR and ST and a mild increase in SC latency. For stress, not at all stressful odors induced an increase in SC amplitude, latency and rise time. Mildly stressful odors induced a very small increase in SC amplitude, latency and rise time, and very stressful odors decreased FP activity, changed in both directions AR activity, increase mildly SC amplitude and induced an increase in SC number of events. Finally, for anxiety, not at all anxious-inducing odors decreased ST and increased SC activity (amplitude, rise time and to a lesser degree latency). Mildly anxious-inducing odors increased very slightly SC amplitude and rise time. Very anxious-inducing odors induced a change in both directions of AR, increased slightly SC amplitude and increased strongly SC latency.

4.3 Multidimensional descriptive rules

The rules with multiple intervals are depicted in Fig. S3 (Supplementary file). Each interval is represented by a colored bar on the axis of the corresponding physiological attributes. A first result of interest is that 92.3% of these multidimensional rules concern the "medium" class (Table 2) of four perceptual dimensions (Pleasantness, Intensity, Stress and Anxiety). The relaxation dimension did not contain any *multidimensional descriptive rules* and therefore contains only simple rules. A second result of interest is that none of the extreme ratings ("low", "high") for the dimensions of intensity, relaxation, stress and anxiety was described by any multiple physiological rules. The only exception was the state "High" pleasantness which was characterized by 20% of multiple rules. In summary, these data suggest that simple physiological rules are in place when the olfactory system is confronted with chemicals evaluated with extreme subjective hedonic scores: odors evaluated as "high" and "low" in terms of pleasantness were associated with no or very few multiple rules. However, it should be kept in mind that these descriptive rules are not systematically the same across individuals. This variability is discussed in Section 4.4.

4.4 Variability in the generated descriptive rules

The number of individuals participating in a given pattern was calculated from the *informedness* of each individual for that pattern. A first visualization in a standardized frame, allowed us to visualize which individuals participate in the pattern and how much they participate to the pattern. The graph of patterns classified as TOP 1 can be viewed in Fig. S4 (Supplementary file). One such example is displayed in Fig. 7 whereby each point depicts a participant labelled with his/her number. The two axes of the figure

correspond to the false positive and true positive rate. Note that *informedness* is the number of true positive ratio minus the false positive ratio and it must be positive for the individual to be noted as a participant in the pattern. In sum, each subject positioned above the unit slope line participates to the pattern (symbolized in white). The more a white point is distant from the unit slope line, the greater the number of true positives compared to false positives and therefore the more the subject participates in the pattern.

A second visualization was carried out in the form of a Venn diagram (see Fig. S5 and S6, Supplementary file) created with Interativenn [72]. Fig. 8a illustrates such data for pleasantness ratings. For each target, we selected the same patterns as those visualized in Fig. 6 and calculated the overlap in terms of individuals for each pattern of the same class. The Venn diagrams thus allow us to illustrate whether it is the same individuals who participate in the different patterns or not. For example, in the 15 individuals who participate in the "1. SC_Latency" pattern of the unpleasant in fig. 8a: 13 are also present in the other two patterns, 1 individual is common only with the pattern "4. Sc_Rise_Time" and 1 individual is not present in the other patterns. In order to be able to compare these 3 diagrams, we reconstructed these diagrams retaining only physiological attributes common to the unpleasant, neutral and pleasant odors (Fig. 8b). An important result to extract from this analysis is that the more unpleasant the subjective emotional experience, the greater the number of individuals participating in the physiological pattern. In other words, the physiological response to aversive stimuli appears to be more invariant than the physiological response to appetitive stimuli.

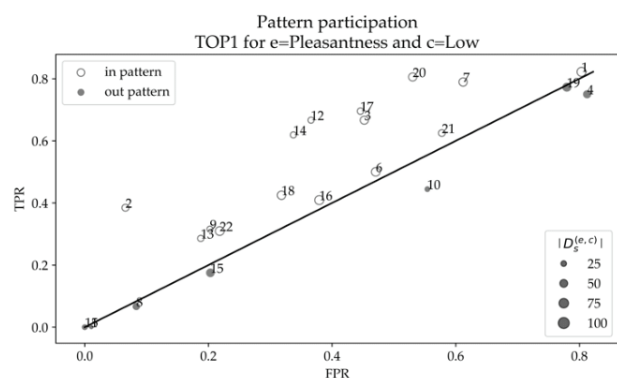


Fig. 7. Participation of each subject in the best pattern found for unpleasant odors. The number displayed on each point corresponds to the subject's identification number. The size of the dot indicates the number of odors classified as unpleasant to that individual.

5 DISCUSSION

The present study was aimed at exploring the physiological underpinnings of odor pleasantness and odor intensity by considering inter-individual diversity. To this end, we used a computational exploratory approach and generated a series of descriptive rules for different subjective experiences. A first result of interest was that our approach provided a set of descriptive rules linking the dimensions of intensity and pleasantness with a number of physiological parameters. For example, for the dimension of olfactory intensity, we observed a gradual increase in the electrodermal response as a function of an increase in intensity. For odor pleasantness, while the perception of pleasant odors is associated with a decrease in cardiac activity, the perception of unpleasant odors is associated with a strong activation of the electrodermal response. These results are in line with previous studies [2], [3], [4], [18], [20] and support the hypothesis of a stronger somatic response for aversive (unpleasant and intense) olfactory stimuli.

A second result of interest concerns the invariance or diversity that characterizes physiological responses to both pleasant and unpleasant stimuli. While for the former we observed great diversity among individuals, the pattern

observed for the latter was very consistent among participants. These data bring a new element of information to the study of subjective hedonic experience of smells by opening the door to the existence of universal somatic and physiological responses between individuals, responses that provide relevant information about our harmful environment [6]. On the other hand, beyond these responses to aversive stimuli, there would exist physiological patterns dedicated to appetitive stimuli, which are important for our well-being, and which would be more sensitive to individual development and trajectory of each and everyone. These results of greater inter-individual variability observed in responses to pleasant and neutral odors compared to unpleasant odors are in agreement with the study by Kroupi et al [61]. In the latter, a classifier discriminated with a high level of prediction the unpleasantness of an odor from EEG data. However, the algorithm struggled to predict the pleasantness and/or neutrality of an odor. Because the explanatory power of such a predictive approach is relatively poor, the explanation for these classification difficulties was not explicit. In another study [60], the authors improved the quality of the model by making predictions within a subject rather than from all subjects' data.

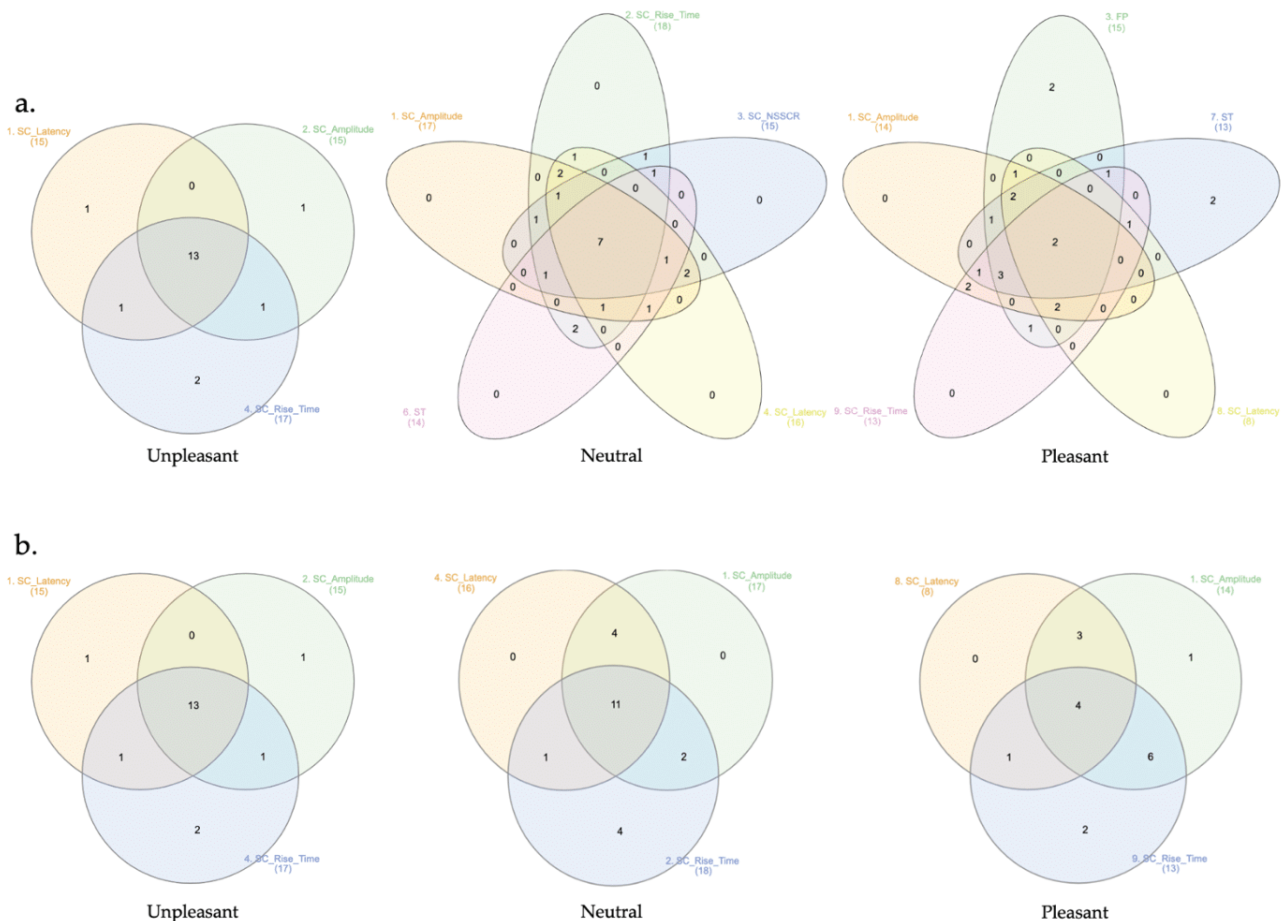


Fig. 8. Venn diagrams of the overlap of individuals in the different physiological patterns for each class of pleasantness by considering all physiological parameters (a) or only those common to the three classes (b). Each circle or ellipse corresponds to a pattern with a simple descriptive rule, it is labeled by its classification number in the TOP10 pattern followed by the physiological attribute concerned by the rule and then by the number of subjects participating in this pattern. The numbers in each diagram indicate both the number of subjects in each intersection and the relative complement in order to know the number of common participants in the overlapping patterns.

This reinforces the idea that inter-individual variability hinders the correct classification of emotions at the population level and that it is important to put the individual at the center of the analysis.

As a secondary objective, we asked the question of whether subjective dimensions such as perceived and reported stress, relaxation or anxiety could have a physiological correlate in the manner already studied for the dimensions of intensity and especially pleasantness. The rules generated for these dimensions do not seem to fit the hypothesis of a decrease in sympathetic activity (decrease in AR, FP, SC and increase in ST) in response to relaxing, non-stressful and non-anxiety-inducing stimuli. This observation leads us to make different hypotheses. The first is that the generated rules do not correspond to a physiological reality and that they depend on the data set or are generated at random. The second is that not all smells have the potential of inducing strong states of anxiety, stress and/or relaxation, and therefore the data distribution within each of these dimensions was not ideal to perform our analysis. Indeed, it may be the case that most of the participants declared to perceive most odorants as not at all (or very) relaxing, stressful and anxiety-inducing stimuli leading to strong disparities in the classes low, medium and high for these dimensions. The third hypothesis is that the participants from the current study exhibited difficulties in understanding the meaning of these dimensions of relaxation, stress and anxiety and thus/or showed difficulties in describing their own feelings about these emotional dimensions when applied to olfaction. Even if it is difficult to provide a definitive answer regarding the acceptance or rejection of one of these hypotheses, we think that the first hypothesis is the less plausible since if it was validated, this would have to apply to all subjective judgements including judgements of pleasantness and intensity. The results for these dimensions of pleasantness and intensity are - however - in good agreement with the literature. Moreover, the second hypothesis is strengthened by the fact that the 3 classes of these 3 affective dimensions were very clearly unbalanced, the rating "1" (not at all) was used widely more (between 33.5 and 45.7%) than the others (see Fig. S7, Supplementary file) (aggregate means and standard deviations in a scale from 1 to 9 across all subjects and all stimuli: 3.36+/-2.32 for relaxation, 2.87+/-2.32 for stress, 2.61+/-2.00 for anxiety). As such, our study cannot conclude whether the second hypothesis and/or the third hypothesis (misunderstanding of the dimensions of relaxation, stress, and anxiety as applied to the universe of smells) is the most plausible. In both cases, and to clarify this limitation, it will be important for future studies in this area to consider i/ the ability of odors to induce these subjective experiences in humans, and ii/ the development of appropriate measurement scales to assess the subjective nature of these dimensions. In particular, it would be relevant to combine within the same scale, different states that may belong to the same subjective dimension. This is the case of the state of relaxation which can be opposed on the same continuum to perceived stress. It is possible that with such a scale, participants can describe their subjective experience more easily. Nevertheless, we note that the results

observed for these dimensions do not compromise our approach in any way, since it proved to be effective for our dimensions of interest, namely pleasantness and intensity.

When such a scientific approach of linking perceptual dimensions (pleasantness and intensity) with a physiological space is undertaken, one question that may be asked is how the data mining approach is positioned in relation to other approaches in artificial intelligence such as machine learning. In fact, these two approaches complement each other. Whereas machine learning allows us to validate predictive models, the data mining approach is more focused on knowledge extraction by providing descriptive and especially explanatory modeling. When datasets are characterized by a large inter-individual variability (which is the case of our data), it sometimes happens that the machine learning approach does not allow to put forward a solid predictive model. To illustrate this with our data, we tested a series of classifiers (e.g., Decision trees (DT), Random Forest (RF), Logistic regression (Log. Reg.), K-nearest neighbors (KNN), Support vector machine (SVM), Multi-layer perceptron (MLP)), all available in the scikit-learn Python library [73] using tuning in k-fold cross-validation. Note that the higher the value of k, the higher the accuracy should be, and it is common to use k-values of 5 or 10. Here, the performance was very similar to different k-values (3, 5 and 10; accuracy difference less than 2%), and a value of k=10 was chosen. The tuning allows to choose a set of optimal hyperparameters for the learning algorithm. The mean cross-validated accuracies of the best estimators are shown in Table 3 for pleasantness and intensity; the other affective dimensions and details on the parameters selected by tuning are available in Table S6 (Supplementary file). Note that the accuracy is never better than the size of the largest class. Indeed, for relaxation, stress and anxiety, DT, RF, Log. Reg, SVM and MLP classify in more than 90% of cases the most frequent class. As mentioned in the Section 2 (Related works), the classification scores are not high suggesting that these different algorithms have not been able to manage the diversity of human perception inherent in this dataset. In this case, the data mining approach can be interesting to implement because it allows to extract a certain number of rules linking olfactory perception and physiology and in fine to open the door to the validation of new hypotheses that we would not necessarily have thought of when analyzing the data manually. It should also be noted that this data mining approach has different objectives than machine learning methods such as feature extraction. Whereas feature selection allows to extract the most relevant predictive variables in order to define the most robust predictive model, in data mining, the extraction is not about feature extraction but about the rules that will link features from two different spaces. In data mining, the generated knowledge is characterized by a high level of interpretability. In our study, these rules are computed by mathematically well-funded measures and consist in conjunctions of conditions on physiological attributes that conclude on particular subjective experiences of odors. The generated rules are easy to assimilate for a domain expert such as psychologists or neuroscientists in our case. Because it enables modeling that extracts

TABLE 3
MEAN 5-FOLD CROSS-VALIDATION ACCURACY (%) OF THE BEST PREDICTORS FOUND BY TUNING

	DT	RF	Log. Reg.	KNN	SVM	MLP	Size of the largest class (%)
Pleasantness	39.3	39.8	39.5	34.4	40.3	41.2	40.1
Intensity	38.7	40.7	40.6	38	41.5	40.8	38.2

descriptive rules from the data that link subgroups belonging to both perceptual and physiological spaces, such approach is positioned upstream of predictive modeling and therefore allows developing new hypotheses in the field. In sum, the rules generated may allow scientists to start a hypothetical-deductive approach, to formulate new scientific assumptions, then to set up new experiments and lastly to test new predictive models using machine learning approaches.

Although the present study provides interesting results linking physiology and odor pleasantness and intensity, some limitations are to be noted and should be discussed. A first limitation concerns the size of the sample studied ($n=22$). Is it sufficient to extract knowledge linking odor pleasantness or odor intensity with physiological responses with such a number of participants? When using conventional statistics, power analysis is used to answer this question. For example, in Licon's original study with these 22 subjects, a significant relationship was observed between physiological responses to odors and odor pleasantness (Licon et al., 2018, page 5). The statistical power associated with this correlation was adequate and exceeded the 80% threshold (83% in this case). In our case, the minimum number of volunteers is actually the one that allows us to observe exceptional patterns that stand out from chance. Even though our analysis revealed a number of exceptional patterns, the fact that our findings are limited by our sample of participants cannot be ruled out. It will therefore be important in future studies to investigate a method for assessing a minimum sample size for such data mining approaches. For example, one could imagine data mining algorithms, that would discover rules on the basis of a first sample and that would set out to verify if these rules are still observed on another independent sample. These future algorithms should also be able to account for the variation that could exist between the rules generated by the 2 datasets (for instance using Jaccard-based measure [74]), and evaluate the optimal sample size capable of generating stable rules. Nevertheless, it should be kept in mind that a recurrent problem in this topic of the relations between perception and physiology in the field of odors concerns the data sets which are very few. The vast majority of studies use a limited number of odorants, unlike our study which is one of the few to include more than 100 molecules. Thus, beyond the development of new data mining algorithms, it will be necessary to accumulate more perceptual and physiological data sets by integrating more olfactory stimuli in the experimental designs.

A second point of improvement concerns the scales used to evaluate the different perceptual and affective dimensions. As mentioned above in this discussion, one can

wonder whether the fact that participants over-represent or under-represent certain categories of the scales may have introduced a bias in our results. Although these patterns are very unfrequent for the main dimensions of pleasantness and intensity (see Figures S1 and S7), it happens for the dimensions of stress, relaxation and anxiety that participants do not have the same distribution of the "low", "medium", "high" categories. For these three dimensions, it is therefore possible that this non-uniform distribution may introduce bias into our data. Note however that subgroup discovery analysis considers in a way this bias by extracting invariant descriptions through both the score and the number of individuals who participate to the pattern. Nevertheless, it will be important in future work to further limit the polarized reactions of the participants by taking into account the intrinsic distribution of each individual by weighting individual's trials according to this distribution. This could be done by adding metrics such as subjective interestingness [75]. A third limitation of our study concerns other physiological parameters that can undoubtedly help to improve knowledge extraction. This will imply not only new signals and modalities, but also more complex features of selected physiological signals. Our protocol focused on physiological signals that were simple enough to measure and already collected in a previous study [17]. A possible perspective of our study would be to complete these measurements by parameters even more discriminating in the classification of affective responses. One can think for example of heart rate variability which can be measured using ECG sensors. The measurement of EEG is also an avenue to follow by considering signal amplitudes and latencies but also more complex features in the time-frequency domain for example. Finally, measurement of brain activity via fMRI should also be considered. For the latter, the problem will be the number of trials per hedonic condition (e.g. pleasant vs. unpleasant) which may be limited due to the temporal resolution of the method (which limits the study in time) and the habituation phenomena (which increases the duration of experiment). Taken as a whole, such an approach will enrich our knowledge of the perceived pleasantness and intensity of smells at different levels of processing (behavior, central nervous system, peripheral nervous system), a multimodal approach that is too rarely implemented.

6 CONCLUSIONS

The main aim of the present study was to set out to examine whether the use of subgroup discovery may help in improving our understanding of the relationship between perceptual dimensions of pleasantness and intensity of

odors and their physiological underpinnings. The presented workflow enabled to relate autonomic nervous system activity patterns with specific odor perceptual responses. On specificity of this method is that it applies to a large amount of data, which is not always the case with all datasets from affective sciences. To circumvent this issue, and generate pertinent hypotheses in the field, one could combine different datasets from different studies with similar protocols since our approach can be used on experiments whose stimuli (olfactory and non-olfactory) are not necessarily the same for all participants. Last but not least, this approach can be applied to other fields of affective sciences characterized by large, complex and heterogeneous datasets.

ACKNOWLEDGMENT

This study was granted by the ANR ChemoSim project.

REFERENCES

- [1] Y. Yeshurun and N. Sobel, "An Odor is Not Worth a Thousand Words: From Multidimensional Odors to Unidimensional Odor Objects," *Annu. Rev. Psychol.*, vol. 61, no. 1, pp. 219–241, Jan. 2010, doi: 10.1146/annurev.psych.60.110707.163639.
- [2] J. Djordjevic, J. N. Lundstrom, F. Clément, J. A. Boyle, S. Pouliot, and M. Jones-Gotman, "A Rose by Any Other Name: Would it Smell as Sweet?," *J. Neurophysiol.*, vol. 99, no. 1, pp. 386–393, Jan. 2008, doi: 10.1152/jn.00896.2007.
- [3] M. Bensafi, C. Rouby, V. Farget, B. Bertrand, M. Vigouroux, and A. Holley, "Psychophysiological correlates of affects in human olfaction," *Neurophysiol. Clin. Neurophysiol.*, vol. 32, no. 5, pp. 326–332, Nov. 2002, doi: 10.1016/S0987-7053(02)00339-8.
- [4] O. Robin, O. Alaoui-Ismaïli, A. Dittmar, and E. Vernet-Maury, "Basic Emotions Evoked by Eugenol Odor Differ According to the Dental Experience. A Neurovegetative Analysis," *Chem. Senses*, vol. 24, no. 3, pp. 327–335, Jan. 1999, doi: 10.1093/chemse/24.3.327.
- [5] C. Ferdenzi *et al.*, "Variability of Affective Responses to Odors: Culture, Gender, and Olfactory Knowledge," *Chem. Senses*, vol. 38, no. 2, pp. 175–186, Feb. 2013, doi: 10.1093/chemse/bjs083.
- [6] P. Jossain, C. Ferdenzi, J. Djordjevic, and M. Bensafi, "Relationship Between Psychophysiological Responses to Aversive Odors and Nutritional Status During Normal Aging," *Chem. Senses*, vol. 42, no. 6, pp. 465–472, Jul. 2017, doi: 10.1093/chemse/bjx027.
- [7] R. J. Davidson, "Anterior cerebral asymmetry and the nature of emotion," *Brain Cogn.*, vol. 20, no. 1, pp. 125–151, Sep. 1992, doi: 10.1016/0278-2626(92)90065-T.
- [8] P. J. Lang, M. K. Greenwald, M. M. Bradley, and A. O. Hamm, "Looking at pictures: Affective, facial, visceral, and behavioral reactions," *Psychophysiology*, vol. 30, no. 3, pp. 261–273, 1993, doi: 10.1111/j.1469-8986.1993.tb03352.x.
- [9] A. J. Gerber *et al.*, "An affective circumplex model of neural systems subserving valence, arousal, and cognitive overlay during the appraisal of emotional faces," *Neuropsychologia*, vol. 46, no. 8, pp. 2129–2139, Jul. 2008, doi: 10.1016/j.neuropsychologia.2008.02.032.
- [10] P. Ekman, "Are there basic emotions?," *Psychol. Rev.*, vol. 99, no. 3, pp. 550–553, 1992, doi: 10.1037/0033-295X.99.3.550.
- [11] R. J. Dolan, "Emotion, Cognition, and Behavior," *Science*, vol. 298, no. 5596, pp. 1191–1194, Nov. 2002, doi: 10.1126/science.1076358.
- [12] E. T. Rolls, "Taste and smell processing in the brain," in *Handbook of Clinical Neurology*, vol. 164, Elsevier, 2019, pp. 97–118. doi: 10.1016/B978-0-444-63855-7.00007-1.
- [13] O. Alaoui-Ismaïli, O. Robin, H. Rada, A. Dittmar, and E. Vernet-Maury, "Basic Emotions Evoked by Odorants: Comparison Between Autonomic Responses and Self-Evaluation," *Physiol. Behav.*, vol. 62, no. 4, pp. 713–720, Oct. 1997, doi: 10.1016/S0031-9384(97)90016-0.
- [14] O. Alaoui-Ismaïli, E. Vernet-Maury, A. Dittmar, G. Delhomme, and J. Chanel, "Odor Hedonics: Connection With Emotional Response Estimated by Autonomic Parameters," *Chem. Senses*, vol. 22, no. 3, pp. 237–248, 1997, doi: 10.1093/chemse/22.3.237.
- [15] E. Vernet-Maury, O. Alaoui-Ismaïli, A. Dittmar, G. Delhomme, and J. Chanel, "Basic emotions induced by odorants: a new approach based on autonomic pattern results," *J. Auton. Nerv. Syst.*, vol. 75, no. 2, pp. 176–183, Feb. 1999, doi: 10.1016/S0165-1838(98)00168-4.
- [16] R. M. Khan *et al.*, "Predicting Odor Pleasantness from Odorant Structure: Pleasantness as a Reflection of the Physical World," *J. Neurosci.*, vol. 27, no. 37, pp. 10015–10023, Sep. 2007, doi: 10.1523/JNEUROSCI.1158-07.2007.
- [17] C. C. Licon, C. Manesse, M. Dantec, A. Fournel, and M. Bensafi, "Pleasantness and trigeminal sensations as salient dimensions in organizing the semantic and physiological spaces of odors," *Sci. Rep.*, vol. 8, no. 1, Art. no. 1, May 2018, doi: 10.1038/s41598-018-26510-5.
- [18] M. Bensafi, C. Rouby, V. Farget, B. Bertrand, M. Vigouroux, and A. Holley, "Autonomic Nervous System Responses to Odours: the Role of Pleasantness and Arousal," *Chem. Senses*, vol. 27, no. 8, pp. 703–709, Oct. 2002, doi: 10.1093/chemse/27.8.703.
- [19] P. Brauchli, P. B. Rüegg, F. Etzweiler, and H. Zeier, "Electrocortical and Autonomic Alteration by Administration of a Pleasant and an Unpleasant Odor," *Chem. Senses*, vol. 20, no. 5, pp. 505–515, 1995, doi: 10.1093/chemse/20.5.505.
- [20] P. Møller and G. Dijksterhuis, "Differential human electrodermal responses to odours," *Neurosci. Lett.*, vol. 346, no. 3, pp. 129–132, Aug. 2003, doi: 10.1016/S0304-3940(03)00498-1.
- [21] M. Nardelli, G. Valenza, A. Greco, A. Lanatá, E. P. Scilingo, and R. Bailón, "Quantifying the lagged Poincaré plot geometry of ultrashort heart rate variability series: automatic recognition of odor hedonic tone," *Med. Biol. Eng. Comput.*, vol. 58, no. 5, pp. 1099–1112, May 2020, doi: 10.1007/s11517-019-02095-7.
- [22] M. Bensafi, C. Rouby, V. Farget, B. Bertrand, M. Vigouroux, and A. Holley, "Influence of affective and cognitive judgments on autonomic parameters during inhalation of pleasant and unpleasant odors in humans," *Neurosci. Lett.*, vol. 319, no. 3, pp. 162–166, Feb. 2002, doi: 10.1016/S0304-3940(01)02572-1.
- [23] M. Mantel, J.-M. Roy, and M. Bensafi, "Accounting

- For Subjectivity In Experimental Research On Human Olfaction,” *Chem. Senses*, no. bjaa082, Jan. 2021, doi: 10.1093/chemse/bjaa082.
- [24] C. Rouby, S. Pouliot, and M. Bensafi, “Odor hedonics and their modulators,” *Food Qual. Prefer.*, vol. 20, no. 8, pp. 545–549, Dec. 2009, doi: 10.1016/j.foodqual.2009.05.004.
- [25] R. Doty and V. Kamath, “The influences of age on olfaction: a review,” *Front. Psychol.*, vol. 5, 2014, Accessed: Apr. 15, 2022. [Online]. Available: <https://www.frontiersin.org/article/10.3389/fpsyg.2014.00020>
- [26] P. Sorokowski *et al.*, “Sex Differences in Human Olfaction: A Meta-Analysis,” *Front. Psychol.*, vol. 10, 2019, Accessed: Apr. 15, 2022. [Online]. Available: <https://www.frontiersin.org/article/10.3389/fpsyg.2019.00242>
- [27] D. Angluin and P. Laird, “Learning from noisy examples,” *Mach. Learn.*, vol. 2, no. 4, pp. 343–370, 1988.
- [28] A. Belfodil *et al.*, “FSSD - A Fast and Efficient Algorithm for Subgroup Set Discovery,” Washington DC, United States, Oct. 2019. Accessed: Jan. 07, 2022. [Online]. Available: <https://hal.archives-ouvertes.fr/hal-02355503>
- [29] R. Mathonat, “Rule discovery in labeled sequential data: application to game analytics.” University of Lyon, France, 2020.
- [30] C. Collet, E. Vernet-Maury, G. Delhomme, and A. Dittmar, “Autonomic nervous system response patterns specificity to basic emotions,” *J. Auton. Nerv. Syst.*, vol. 62, no. 1, pp. 45–57, Jan. 1997, doi: 10.1016/S0165-1838(96)00108-7.
- [31] W. He, S. Boesveldt, C. de Graaf, and R. A. de Wijk, “Dynamics of autonomic nervous system responses and facial expressions to odors,” *Front. Psychol.*, vol. 5, 2014, doi: 10.3389/fpsyg.2014.00110.
- [32] S. A. Neumann and S. R. Waldstein, “Similar patterns of cardiovascular response during emotional activation as a function of affective valence and arousal and gender,” *J. Psychosom. Res.*, vol. 50, no. 5, pp. 245–253, May 2001, doi: 10.1016/S0022-3999(01)00198-2.
- [33] D. Palomba, M. Sarlo, A. Angrilli, A. Mini, and L. Stegagno, “Cardiac responses associated with affective processing of unpleasant film stimuli,” *Int. J. Psychophysiol.*, vol. 36, no. 1, pp. 45–57, Apr. 2000, doi: 10.1016/S0167-8760(99)00099-9.
- [34] T. Ritz, M. Thöns, S. Fahrenkrug, and B. Dahme, “Airways, respiration, and respiratory sinus arrhythmia during picture viewing,” *Psychophysiology*, vol. 42, no. 5, pp. 568–578, 2005, doi: 10.1111/j.1469-8986.2005.00312.x.
- [35] E. P. M. Vianna and D. Tranel, “Gastric myoelectrical activity as an index of emotional arousal,” *Int. J. Psychophysiol.*, vol. 61, no. 1, pp. 70–76, Jul. 2006, doi: 10.1016/j.ijpsycho.2005.10.019.
- [36] S. R. Vrana, “The psychophysiology of disgust: Differentiating negative emotional contexts with facial EMG,” *Psychophysiology*, vol. 30, no. 3, pp. 279–286, 1993, doi: 10.1111/j.1469-8986.1993.tb03354.x.
- [37] R. A. de Wijk, V. Kooijman, R. H. G. Verhoeven, N. T. E. Holthuysen, and C. de Graaf, “Autonomic nervous system responses on and facial expressions to the sight, smell, and taste of liked and disliked foods,” *Food Qual. Prefer.*, vol. 26, no. 2, pp. 196–203, Dec. 2012, doi: 10.1016/j.foodqual.2012.04.015.
- [38] Paul Ekman and R. W. Levenson, “Autonomic nervous system activity distinguishes between emotions. *Science*,” *Science*, vol. 221, no. 86, pp. 1208–1210, 1983.
- [39] R. W. Levenson, K. Heider, and P. Ekman, “Emotion and Autonomic Nervous System Activity in the Minangkabau of West Sumatra,” p. 17.
- [40] M. Shiota, S. Neufeld, W. Yeung, S. Moser, and E. Perea, “Feeling Good: Autonomic Nervous System Responding in Five Positive Emotions,” *Emot. Wash. DC*, vol. 11, pp. 1368–78, Dec. 2011, doi: 10.1037/a0024278.
- [41] I. Van Diest *et al.*, “Critical Conditions for Hyperventilation Responses: The Role of Autonomic Response Propositions During Emotional Imagery,” *Behav. Modif.*, vol. 25, no. 4, pp. 621–639, Sep. 2001, doi: 10.1177/0145445501254008.
- [42] U. Hess, A. Kappas, G. J. McHugo, J. T. Lanzetta, and R. E. Kleck, “The facilitative effect of facial expression on the self-generation of emotion,” *Int. J. Psychophysiol.*, vol. 12, no. 3, pp. 251–265, May 1992, doi: 10.1016/0167-8760(92)90064-I.
- [43] S. Delplanque *et al.*, “Sequential unfolding of novelty and pleasantness appraisals of odors: Evidence from facial electromyography and autonomic reactions,” *Emotion*, vol. 9, no. 3, pp. 316–328, 2009, doi: 10.1037/a0015369.
- [44] S. Watanuki and Y.-K. Kim, “Physiological Responses Induced by Pleasant Stimuli,” *J. Physiol. Anthropol. Appl. Human Sci.*, vol. 24, no. 1, pp. 135–138, 2005, doi: 10.2114/jpa.24.135.
- [45] C. C. Licon *et al.*, “Chemical features mining provides new descriptive structure-odor relationships,” *PLOS Comput. Biol.*, vol. 15, no. 4, p. e1006945, Apr. 2019, doi: 10.1371/journal.pcbi.1006945.
- [46] B. Sanchez-Lengeling, J. N. Wei, B. K. Lee, R. C. Gerkin, A. Aspuru-Guzik, and A. B. Wiltschko, “Machine Learning for Scent: Learning Generalizable Perceptual Representations of Small Molecules,” *ArXiv191010685 Phys. Stat.*, Oct. 2019, Accessed: Jul. 13, 2021. [Online]. Available: <http://arxiv.org/abs/1910.10685>
- [47] A. Choi and W. Woo, “Physiological Sensing and Feature Extraction for Emotion Recognition by Exploiting Acupuncture Spots,” in *Affective Computing and Intelligent Interaction*, Berlin, Heidelberg, 2005, pp. 590–597. doi: 10.1007/11573548_76.
- [48] A. Haag, S. Goronzy, P. Schaich, and J. Williams, “Emotion Recognition Using Bio-sensors: First Steps towards an Automatic System,” in *Affective Dialogue Systems*, Berlin, Heidelberg, 2004, pp. 36–48. doi: 10.1007/978-3-540-24842-2_4.
- [49] C. D. Katsis, N. Katertsidis, G. Ganiatsas, and D. I. Fotiadis, “Toward Emotion Recognition in Car-Racing Drivers: A Biosignal Processing Approach,” *IEEE Trans. Syst. Man Cybern. - Part Syst. Hum.*, vol. 38, no. 3, pp. 502–512, May 2008, doi: 10.1109/TSMCA.2008.918624.
- [50] C. D. Katsis, N. S. Katertsidis, and D. I. Fotiadis, “An integrated system based on physiological signals for the assessment of affective states in patients with anxiety disorders,” *Biomed. Signal Process. Control*, vol. 6, no. 3, pp. 261–268, Jul. 2011, doi: 10.1016/j.bspc.2010.12.001.
- [51] C. L. Lisetti and F. Nasoz, “Using Noninvasive Wearable Computers to Recognize Human Emotions from Physiological Signals,” *EURASIP J. Adv. Signal Process.*, vol. 2004, no. 11, p. 929414, Dec. 2004, doi: 10.1155/S1110865704406192.
- [52] S. Yoo, C. Lee, Y. Park, N. Kim, B. Lee, and K. Jeong, *Neural Network Based Emotion Estimation Using Heart Rate Variability and Skin Resistance*, vol. 3610. 2005, p. 824. doi:

10.1007/11539087_110.

- [53] P. A. Kragel and K. S. LaBar, "Multivariate Pattern Classification Reveals Autonomic and Experiential Representations of Discrete Emotions," *Emot. Wash. DC*, vol. 13, no. 4, pp. 681–690, Aug. 2013, doi: 10.1037/a0031820.
- [54] C. Liu, K. Conn, N. Sarkar, and W. Stone, "Physiology-based affect recognition for computer-assisted intervention of children with Autism Spectrum Disorder," *Int. J. Hum.-Comput. Stud.*, vol. 66, no. 9, pp. 662–677, Sep. 2008, doi: 10.1016/j.ijhsc.2008.04.003.
- [55] J. Zhai and A. Barreto, "Stress Detection in Computer Users Based on Digital Signal Processing of Noninvasive Physiological Variables," in *2006 International Conference of the IEEE Engineering in Medicine and Biology Society*, Aug. 2006, pp. 1355–1358. doi: 10.1109/IEMBS.2006.259421.
- [56] C. L. Stephens, I. C. Christie, and B. H. Friedman, "Autonomic specificity of basic emotions: Evidence from pattern classification and cluster analysis," *Biol. Psychol.*, vol. 84, no. 3, pp. 463–473, Jul. 2010, doi: 10.1016/j.biopsycho.2010.03.014.
- [57] I. C. Christie and B. H. Friedman, "Autonomic specificity of discrete emotion and dimensions of affective space: a multivariate approach," *Int. J. Psychophysiol.*, vol. 51, no. 2, pp. 143–153, Jan. 2004, doi: 10.1016/j.ijpsycho.2003.08.002.
- [58] C. A. Frantzidis *et al.*, "On the Classification of Emotional Biosignals Evoked While Viewing Affective Pictures: An Integrated Data-Mining-Based Approach for Healthcare Applications," *IEEE Trans. Inf. Technol. Biomed.*, vol. 14, no. 2, pp. 309–318, Mar. 2010, doi: 10.1109/TITB.2009.2038481.
- [59] J. A. Healey and R. W. Picard, "Detecting stress during real-world driving tasks using physiological sensors," *IEEE Trans. Intell. Transp. Syst.*, vol. 6, no. 2, pp. 156–166, Jun. 2005, doi: 10.1109/TITS.2005.848368.
- [60] J. Kim and E. André, "Emotion recognition based on physiological changes in music listening," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 30, no. 12, pp. 2067–2083, Dec. 2008, doi: 10.1109/TPAMI.2008.26.
- [61] E. Kroupi, J. Vesin, and T. Ebrahimi, "Subject-Independent Odor Pleasantness Classification Using Brain and Peripheral Signals," *IEEE Trans. Affect. Comput.*, vol. 7, no. 4, pp. 422–434, Oct. 2016, doi: 10.1109/TAFFC.2015.2496310.
- [62] "Emotion perception and overconfidence in errors under stress in psychosis - ScienceDirect." https://www.sciencedirect.com/science/article/pii/S0165178117310818?casa_token=3TSACKgxaBkAAAAA:dU5e5UnBIKft8DQy-wEx0XFKOm_dF2lKnIAKMOli_PZo-bMi2mvco473Vt7JH6jSEbOdRDx5U6Q (accessed Apr. 19, 2022).
- [63] L. X. Liao, A. M. Corsi, P. Chrysochou, and L. Lockshin, "Emotional responses towards food packaging: A joint application of self-report and physiological measures of emotion," *Food Qual. Prefer.*, vol. 42, pp. 48–55, Jun. 2015, doi: 10.1016/j.foodqual.2015.01.009.
- [64] M. E. Dawson, A. M. Schell, and D. L. Filion, "The electrodermal system," in *Handbook of psychophysiology, 4th ed.*, New York, NY, US: Cambridge University Press, 2017, pp. 217–243.
- [65] J. Macqueen, "Some methods for classification and analysis of multivariate observations," *Proc. Fifth Berkeley Symp. Math. Stat. Probab.*, vol. 1, no. 14, pp. 281–297.
- [66] A. K. Anderson *et al.*, "Dissociated neural representations of intensity and valence in human olfaction," *Nat. Neurosci.*, vol. 6, no. 2, pp. 196–202, Feb. 2003, doi: 10.1038/nn1001.
- [67] H.-R. Hou, X.-N. Zhang, and Q.-H. Meng, "Odor-induced emotion recognition based on average frequency band division of EEG signals," *J. Neurosci. Methods*, vol. 334, p. 108599, Mar. 2020, doi: 10.1016/j.jneumeth.2020.108599.
- [68] J. Fürnkranz, D. Gamberger, and N. Lavrač, *Foundations of Rule Learning*. Springer Science & Business Media, 2012.
- [69] T. Abudawood and P. Flach, "Evaluation Measures for Multi-class Subgroup Discovery," in *Machine Learning and Knowledge Discovery in Databases*, Berlin, Heidelberg, 2009, pp. 35–50. doi: 10.1007/978-3-642-04180-8_20.
- [70] D. M. Powers, "Recall & Precision versus The Bookmaker," *Int. Conf. Cogn. Sci.*, Jul. 2003, Accessed: Feb. 17, 2021. [Online]. Available: <https://dspace.flinders.edu.au/xmlui/handle/2328/27159>
- [71] C. B. Browne *et al.*, "A Survey of Monte Carlo Tree Search Methods," *IEEE Trans. Comput. Intell. AI Games*, vol. 4, no. 1, pp. 1–43, Mar. 2012, doi: 10.1109/TCIAIG.2012.2186810.
- [72] H. Heberle, G. V. Meirelles, F. R. da Silva, G. P. Telles, and R. Minghim, "InteractiVenn: a web-based tool for the analysis of sets through Venn diagrams," *BMC Bioinformatics*, vol. 16, no. 1, p. 169, May 2015, doi: 10.1186/s12859-015-0611-3.
- [73] F. Pedregosa *et al.*, "Scikit-learn: Machine Learning in Python," *J. Mach. Learn. Res.*, vol. 12, pp. 2825–2830, 2011.
- [74] E. Galbrun and P. Miettinen, "What Is Redescription Mining," in *Redescription Mining*, E. Galbrun and P. Miettinen, Eds. Cham: Springer International Publishing, 2017, pp. 1–23. doi: 10.1007/978-3-319-72889-6_1.
- [75] T. De Bie, "Maximum entropy models and subjective interestingness: an application to tiles in binary databases," *Data Min. Knowl. Discov.*, vol. 23, no. 3, pp. 407–446, Nov. 2011, doi: 10.1007/s10618-010-0209-3.



Maëlle Moranges received the B.Sc. in computer science in 2016 and the M.Sc. degree in bioinformatics in 2018 from University Lyon1. She is currently working toward the Ph.D. degree at Lyon Neuroscience Research Center (France). Her research interests include the study of emotional variability and neurophysiological variability in olfactory perception.



Marc Plantevit received the PhD degree in computer science from the University of Montpellier in 2008. He is a full professor at EPITA, and member of the EPITA research lab. Before joining EPITA, he was an associate professor at University Claude Bernard Lyon 1. His research is mainly concerned with foundation of data mining, graph mining, subgroup discovery and explainable artificial intelligence. He is on the editorial board of Data Mining and Knowledge Discovery Journal.



Moustafa Bensafi received a PhD in Cognitive Psychology and an HDR in Neuroscience at the University of Lyon. He is currently CNRS Research Director at Lyon Neuroscience Research Center (LNRC, France). He is a co-leader of the Neuropop team at LNRC. His research is focusing on the behavioral and neural bases of chemosensory processing in humans.

ÉTUDE 2 : « Utilisation de l'extraction de sous-graphes attribués exceptionnels pour explorer la variabilité interindividuelle dans le traitement du caractère hédonique des odeurs au sein du cortex piriforme et l'amygdale»
(en préparation)

Matériel supplémentaire : [Annexe 5](#)

Using exceptional attributed subgraph mining to explore interindividual variability in odor pleasantness processing in the human piriform cortex and amygdala

Maëlle Moranges^{1,2}, Arnaud Fournel¹, Marc Thévenet¹, Marc Plantevit^{2,3}, Moustafa Bensafi¹

¹Lyon Neuroscience Research Center, CNRS, INSERM, Université Lyon1

²Laboratoire d'InfoRmatique en Image et Systèmes d'information, Université Lyon1

³EPITA Research and Development Laboratory (LRDE), France

Acknowledgements: We thank Frédéric Faure, Pauline Jousain and Caroline Sezille from the University of Lyon 1, and Danielle Ibarrola and her team at the CERMEP brain imaging center.

Funding: This research was funded by the CNRS and the ANR (ChemoSim project).

Abstract

In humans, the amygdala and piriform cortex are two important brain structures involved in hedonic odor processing. Although the affective processing of odors in these two structures has been extensively studied in the past, the involvement of individual volunteers in global patterns of activation has been little studied. The purpose of this study is to examine whether exceptional pattern extraction techniques can improve our understanding of hedonic odor processing in these brain areas while paying particular attention to individual variability.

A total of 42 volunteers participated in an fMRI study in which they were asked to smell 6 odors and describe their hedonic valence. Classical univariate analyses (SPM type) and data mining were performed on the fMRI data. The results showed convergences between the two analyses: unpleasant odors preferentially activate the anterior part of the left piriform cortex. Moreover, the data mining approach revealed specific patterns for pleasant and unpleasant odors in the piriform cortex but also in the amygdala. This approach also revealed the contribution of each of the 42 individuals to the observed patterns. Taken together, these results suggest that the data mining approach can be used - with standard fMRI analyses - to provide complementary information regarding spatial location and contribution of individuals to the observed pattern.

INTRODUCTION

The sense of smell is an important sensory function for humans: it allows to detect dangers in the environment but also to feel pleasure (1,2). Psychophysical and neurobiological studies showed that this hedonic processing of odors is omnipresent during the cerebral processing of an odor: from a neural point of view, the pleasant/unpleasant character of an odor is represented at several levels from the olfactory epithelium (3), the olfactory bulb (4), the piriform cortex (5,6), the amygdala (7) or the orbitofrontal cortex (8).

Another important aspect of these hedonic processes in relation to odors is that they can be very variable from one individual to another. While for some people the eugenol molecule may remind of the smell of cloves and be perceived as rather pleasant, for others it will remind of the rather unpleasant context of the dentist (9). These inter-individual variations in response to smells are little taken into account in neuroscientific studies that explored the cerebral substrate of hedonic processing of odors (10). Moreover, this variability could explain, at least in part, contradictory results about the role of certain structures in olfactory affective processes from one study to another.

Although some studies show that activity within the piriform cortex and amygdala are not associated with valence but rather with the intensity of the olfactory stimulus (11–13), other studies state that the hedonic valence of odors is represented in both structures (7,14). Still others suggest that the amygdala encodes neither stimulus intensity nor odor valence, but a combination of both dimensions that reflects the overall emotional salience of the olfactory stimulus (15). Furthermore, studies that consider valence independently of intensity show that pleasant and unpleasant odors induce a greater neural response in the amygdala (16) and anterior piriform cortex (14) compared to odors described as neutral (without salient emotional valence). Moreover, it appears that this region encompassing the piriform cortex and the amygdala is more activated for unpleasant odors than for pleasant ones (8), with a preponderant involvement of the anterior portion for the piriform cortex. A lateralization of these olfactory hedonic processes has also been mentioned in the left part of the anterior piriform cortex (5,17). It should be noted that other studies suggest that the posterior portion of the cortex is more involved in processing the quality of an odor (e.g. fruity, floral) than the hedonic valence or the perceived intensity of the stimulus (14,18).

In summary, we can understand from this series of brain imaging studies of the sense of smell that the piriform cortex and the amygdala are indeed involved in odor hedonic processes. Nevertheless, it remains to clarify the spatial localization of pleasant and unpleasant hedonic representations within this piriform/amygdala network, and especially to understand how inter-individual variability in odor hedonic perception can help understanding odor pleasantness processing and brain activity in piriform cortex and amygdala. Although the most commonly used statistical methods (e.g. General Linear Model, Classification, Multivariate analysis etc.) allow a fine analysis of pleasant and unpleasant hedonic representations of odors, they are most often used in group analyses that combine information from all subjects and take little account of inter-individual variability. Interestingly, beyond these conventional approaches mentioned above, there are other techniques from data science that are still too rarely used in the field of brain imaging and that could nevertheless provide complementary information to standard approaches. Among these approaches, one can mention exceptional model mining.

Exceptional model mining is based on the discovery of exceptional subgraph thanks to a *C-energetics* algorithm (19). It detects hotspots for different experimental conditions (ex. pleasant and unpleasant odors) in specific spaces (ex. piriform cortex and/or amygdala). This approach is freed from smoothing thanks to modeling in the form of a graph; the graph allows considering not only one voxel but a voxel and its adjacent voxels. It also overcomes the normalization of the hemodynamic signal per individual by considering the order of activation of experimental trials (ex. odors) for each individual. Interestingly, once the subgraphs have been generated, it is possible to quantify and study the participation of each individual in the pattern. This allows a more fine-grained analysis of the individual contribution of each participant to activation patterns qualified as invariant.

This paper presents an analysis framework on a dataset comprising 42 volunteers who smelled 6 odors and whose hemodynamic activity in response to these stimuli was recorded using fMRI. The data are first processed with conventional univariate analyses (e.g. GLM and region of interest analysis in SPM). In a second step, the data were subjected to Exceptional model mining in order 1/ to provide specific spatial activation patterns in response to the olfactory conditions (e.g. pleasant, unpleasant) to be compared to the results of the univariate analysis and 2/ to provide complementary information to these univariate analyses by modeling the contribution of each participant to the activation spatial patterns observed in each region of interest (e.g. piriform cortex and amygdala) and each olfactory condition (e.g. pleasant, unpleasant).

2. MATERIAL AND METHODS

2.1. Experimental procedure

2.1.1. Participants

The present study consists of a dataset of 42 participants divided into 3 sub-sets. The first sub-data set (subset 1) is original and concerns 14 right-handed young adults (6 males and 8 females; mean age 22.86 ± 3.25 years) who smelled 6 odors varying in hedonic quality. The second dataset (subset 2) concerns 15 right-handed young adults (mean age 22.13 years ± 4.26 years; 7 males and 8 females) who smelled 6 odors also varying in hedonic quality, but different from those used in subset 1. Note that this second dataset has already been exploited in a previous study (20). Finally, the third subset of data (subset 3) is original and involves 13 elderly right-handed individuals (mean age 63.0 years ± 4.08 years; 4 males and 9 females) who smelled the same 6 odors as volunteers of dataset 2. With such an experimental design, it is thus possible to study whether the results observed in elderly vs. young adults, or participants who smelled different sets of odors, participate in the same way in the observed patterns. It should of course be noted that the 42 participants that make up this dataset, whatever the subset, performed the experiment under the same experimental conditions (same MRI scanner, same olfactometer, same experimental task etc.). Only their age or the set of odors differed.

Absence of olfactory deficit was assessed using the European Test of Olfactory Capabilities (ETOC) (21,22), together with a detailed medical history and ENT examination. They received financial compensation for the time spent in the laboratory. The recording

procedure was explained in great detail to the subjects, who provided written consent prior to participation. The study was conducted according to the Declaration of Helsinki and was approved by the Lyon Sud-Est ethics committee.

2.1.2. Odorants

All odorants were provided by Sigma-Aldrich. For subset 1, the six odorants used were (CID, odorant code, vol/vol concentrations): heptanal (HEP; 8130; 0.15%), 1-decanol (8174; DEC; 100%), acetophenone (7410; ACE; 4.8%), eugenol (3314; EUG; 59%), methyl anthranilate (8635; MAN; 65.7%), 3-hexanol (12178; 3HEX; 0.17%). For subset 2, the six odorants were: propanol (1031; PRO; 1.5%), isoamyl acetate (31276; ISO; 0.6%), benzaldehyde (240; BEN; 75%), citronellal (7794; CAL; 15%), citronellol (8842; COL; 22.5%) and trans-2-hexenyl acetate (17243; THA; 1.5%). Note that to ensure iso-intense perception, odorants were individually diluted in mineral oil (see their respective vol/vol concentrations). Whatever the subset, the odorants were delivered by an air-dilution olfactometer, described in detail in Sezille et al. (23). Briefly, the device includes a stimulation system, a mixing head coupled to a delivery system enabling diffusion of odorized air in the subject's nose, a respiratory sensor that triggers the olfactometer according to the subject's nasal respiration, and a response box used to collect subjective odor evaluations. The olfactometer, digitized visual instruction generator, recording of respiratory data, and the MRI scanner itself, were all linked through one TTL (transistor–transistor logic) pulse that assured accurate time-locking of all experimental components. This was done by recording intra-nasal sniffing continuously during the experimental session (see section 2.1.3).

2.1.3. Experimental protocol

An event-related design was used, comprising the odorants (10 trials per stimulus, 20-second inter-stimulus interval, and 5-second stimulus duration) and a non-odorized clean air condition, all trials distributed randomly across 5 fMRI scans (sessions). Each trial began with a 2-sec visual primer “Breathe naturally”, followed 5 seconds later by a 5-sec hedonic task instruction (“Rate the hedonic valence”). Practically, participants were instructed to breathe naturally and after the presentation of each olfactory stimulus (for 5 seconds), they were asked to evaluate odor hedonic valence using a response box with 5 buttons positioned under the 5 fingers of the dominant hand. The participants had 5 options: very unpleasant smell, unpleasant smell, neutral smell, pleasant smell, very pleasant smell. The positioning of the options under the fingers was counterbalanced from one participant to another (e.g. the thumb corresponded to very pleasant for half of the subjects and very unpleasant for the other half). Odorants were diffused synchronously with the subject's nasal respiration: a 5-second stimulus duration was chosen because the odor was released during exhalation and had to be maintained during at least the whole duration of the subsequent inhalation (~2 sec). The recorded signals were: respiratory signal, odor valve opening and TR (time repetition) signal from the fMRI scanner, enabling event-related statistical analysis. Subject's respiratory signal was acquired using an airflow sensor that was integrated on an amplifier interface. A microbridge mass airflow (AWM2100V, Honeywell, MN, USA) allowed acquisition of both inhalation and exhalation phases. The airflow sensor was connected to a nasal cannula (Cardinal Health, OH, USA;

2.8mm inner diameter tube) positioned in both nostrils. Sniffing was digitally recorded at 100 Hz and stored in a computer. Sniffs were pre-processed by removing baseline offsets, and aligned in time by setting the point where the sniff entered the inspiratory phase as time zero. Inspired volume was calculated for the first sniff of every trial and was used as a covariate in the fMRI contrast estimation.

Note that at the end of the fMRI sessions, participants were asked to rate the odorants in terms of intensity, pleasantness, familiarity, edibility, warmth, coolness, irritation and pain, using a visual rating scale ranging from -2 (very unpleasant) to 2 (very pleasant), or from 0 (not at all intense, familiar, edible, warm, cool, painful (pungent) or irritating) to 4 (very intense, familiar, edible, warm, cool, painful (pungent) or irritating).

2.1.4. fMRI parameters

The experiment, which lasted approximately 60 min (from subject's arrival to departure), was performed on a 1.5 Tesla MR-scanner (Siemens Magnetom). The fMRI data were collected in 142 volumes/session (interleaved, AC-AP acquisition) with a 29 axial-slice 2D EPI sequence (matrix: 64x64; TR: 2,500ms; TE: 50ms; FA: 90°; voxel size: 3.43x3.43x3.4mm; FOV: 220). In the 9 minutes immediately following the fMRI session, a high-resolution T1-weighted brain image (3D MPR sequence: TR=1,970ms / TE=3.93ms) was acquired.

2.2. Data analysis

2.2.1. Preprocessing of perceptual data

Hedonic valence scores entered by the participant during the MRI study were discretized into 3 classes: Unpleasant (including the two options of Very unpleasant and unpleasant), Neutral (including the option neutral), and Pleasant (including the two options pleasant and very pleasant). In total, the proportions of responses between the 3 categories across all participants were fairly well balanced (Unpleasant: 35.32%; Neutral: 36.90%; Pleasant: 27.78%). The discretization was performed for each volunteer independently with a K-Means clustering algorithm (24). K-Means divides the subjective ratings into K clusters so that the intra-cluster distance of the scores is minimized and the inter-cluster distance is maximized. This technique is independent of participants' scoring strategies because the discretization is applied at the individual level and yields the 3 categories of odors experienced as differently as possible in terms of pleasantness (25). We chose K=3 so that the cluster with the highest values corresponds to pleasant odors, those with the lowest values correspond to unpleasant odors and the middle cluster to neutral odors.

2.2.2. fMRI Preprocessing and fist level analysis

Functional MRI images were preprocessed using *fMRIPrep* (26) include the following standard steps: 1/ realignment of the voxels to correct the movements, 2/ correction for temporal acquisitions, 3/ co-registration of functional and structural images, 4/ segmentation of the T1 image, 5/ application of the deformation so that the images of

the individual fit into the template image, 6/ estimation of regressors of no interest. At this stage of pre-processing, two datasets have been constructed. For the univariate analyses (section 2.2.3), voxels were smoothed using a 5 x 5 x 5 mm³ FWHM Gaussian kernel to limit potential noise. For the data-mining analyses (section 2.2.4), the data were left in their raw state because the algorithm searches for a set of adjacent voxels having an exceptional and similar behavior. Smoothing the data would artificially create similar behavior between adjacent voxels and the algorithm would thus return biased patterns. The algorithm maximizes the size of the pattern so that a noisy voxel will not stand out in the results. Then, these two datasets (smoothed and unsmoothed) were subjected to first-level analyses. Here, to obtain the estimated activity of each voxel and for each odor, voxel responses were modeled using a design matrix : (i) built with a canonical hemodynamic response function and regressors of interest corresponding to the four conditions (pleasant, neutral, unpleasant, odors) for the univariate analysis, (ii) built with a canonical hemodynamic response function and regressors of interest corresponding to the six odors (Subset 1: HEP, MAN, 3HEX, DEC, ACE, EUG ; Subsets 2 and 3: PRO, ISO, CAL, BEN, COL, THA) for the data mining analysis. The following signals were added to the model as nuisance regressors: 6 first *aCompCor* components (27), 6 motion parameters, frame displacements, nasal respiration (downsampled to fMRI frequency), a cosine basis set acting as a high-pass filter (with a 128s cutoff). Finally, for each participant, the final contrasts consisted in comparing each of the 3 conditions of pleasantness (for the univariate analyses) or the 6 odors (for the data-mining analysis).

2.2.3. Univariate analysis

The univariate analysis is not innovative in the present paper. As its purpose is to present the data using standard approaches that are widely documented in the literature, we will present it in a summary manner and refer the reader to the already existing documentation (see (28)). Here the approach is based on a group analysis on the SPM software which is obtained using single-sample t-tests ($p < 0.001$, uncorrected). The coordinates of the MNI space are used to present all activations. Then, a specific analysis in our different regions of interest (piriform cortex, amygdala) was performed (region of interest analysis or ROI). Note that for the piriform cortex ROI, we distinguished its anterior part from its posterior part since previous studies suggest a different functional role in each of these subregions. The amygdala ROI (both in the left and right hemispheres) was obtained from the automated anatomical atlas (AAL) template (29–31). Anterior and posterior piriform left/right ROIs activity were extracted using hand drawn ROIs. These ROIs were drawn with the MRIcro application using 60 participants who had participated in previous studies (32). Resulting mean images were thresholded to obtain a binary mask: at least 30% of participants had to have a voxel in their ROI, to keep it in the binary mask. We then performed ANOVAs on the average activity extracted in each ROI to examine whether the activation of the regions of interest differ between hemispheres and whether the activation of the different ROIs differentiates pleasant, neutral and unpleasant stimuli. We also looked for an interaction between these two dimensions.

2.2.4. Data mining analysis

The data mining analysis is composed of five steps. The first is common with univariate analysis and concerns preprocessing as was presented previously (2.2.1 and 2.2.2). In this section, we introduce the remaining steps. We will start by modeling the data in the form of an attributed graph, and we will continue with the generation of patterns using data mining algorithms. Next, we will see the validation of these results by bootstrap and finally we will calculate the participation of each individual in the observed patterns. This entire workflow is illustrated in Figure. 1.

Note that in this analysis, our interest is not in the value of the voxels for each odor but rather on the study of the order of the hemodynamic responses of the odors for each voxel by subject. In this way, there is no need to normalize the signal of the participants to compare them.

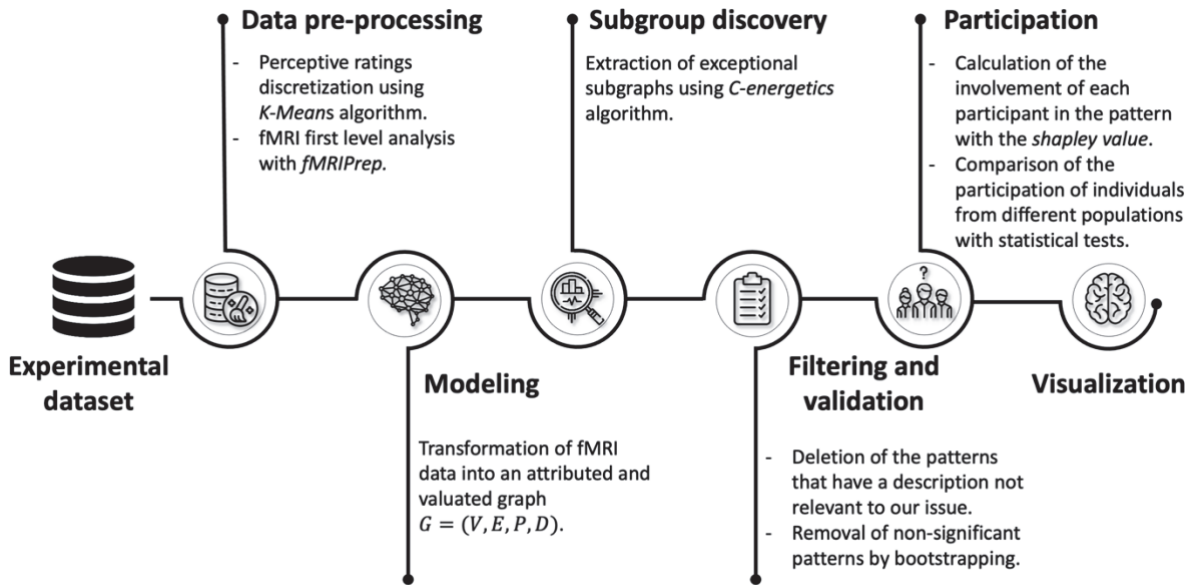


Fig. 1. Steps of the workflow. The raw fMRI data from the experiment is pre-processed and the subjective scores are discretized into 3 categories for each individual. Once the data is cleaned, it is transformed into an attributed and valued graph. This graph is given as input to the subgroup discovery algorithm in order to extract voxel subsets describing the different perceptual dimensions. Then, a statistical validation of the generated patterns is performed: the pattern is validated if its quality measure called WRAcc is outside the confidence interval of the distribution of a bootstrap of 10,000 random draws. Finally, the participation of each individual to the pattern is calculated and the results can be visualized.

i. Modeling

The aim of this first step, is to transform the fMRI data into an attributed graph. These data are composed of one fMRI image by participant (or individual) and by odor and a subjective hedonic rating with this odor. Let $A = \{ 'pleasant', 'neutral', 'unpleasant' \}$ be the set of hedonic categories studied. Let $\pi_i[o]$ such that $\pi_i[o] \in A$ be the category used by the individual i to describe the pleasantness of the odorant $o \in O$. O is the set of odorants such that $O = \{ 'HEP', 'MAN', '3HEX', 'DEC', 'ACE', 'EUG', 'PRO', 'ISO', 'CAL', 'BEN', 'COL', 'THA' \}$.

An fMRI image is composed of several voxels and each voxel contains a hemodynamic activation value β . These voxels are considered here as the vertices V of a graph $G = (V, E, P, D)$ whose edges E connect the adjacent voxels in the fMRI image. The

vertex v representing a voxel at position (x,y,z) is thus connected to the voxels at the following positions: $(x-1,y,z)$, $(x+1,y,z)$, $(x,y-1,z)$, $(x,y+1,z)$, $(x,y,z-1)$ and $(x,y,z+1)$, if these neighboring voxels are included in the region of interest studied.

The set $P = \{(\alpha, \alpha') \mid \alpha \in A, \alpha' \in A \text{ and } \alpha \neq \alpha'\}$ contains all combinations of two non-identical categories. As an example, the pair $(\text{'pleasant'}, \text{'neutral'}) \in P$ aims at depicting the number of cases where odors perceived as neutral have a stronger hemodynamic response than pleasant-perceived odors. To this end, we need to count for each vertex (e.g., voxel) the number of times each possible pair appears.

We denote by $X(v, i, o)$ the level of β activity measured in the vertex v for an individual $i \in [1, n]$ while smelling an odorant o .

We introduce as vertex attributes the pairs of categories from P . Therefore, for each vertex v , each pair $p = (\alpha, \alpha')$ is associated with a unique value d thanks to the function $d(p, v)$:

$$d(p, v) = \sum_{i=1}^n \frac{|\{(o_1, o_2) \in O^2 \mid \pi_i[o_1] = \alpha, \pi_i[o_2] = \alpha', X(v, i, o_1) < X(v, i, o_2)\}|}{|\{o \in O \mid \pi_i[o] = \alpha\}| \cdot |\{o \in O \mid \pi_i[o] = \alpha'\}|}, \text{ with } p = (\alpha, \alpha')$$

The value d for a vertex (i.e., a voxel) v and a pair p is the sum of the average occurrence of the pair for each individual. The average occurrence of a pair for an individual is between 0 and 1; the value d for n individuals is therefore between 0 and n . The sum of the value d for a pair and its opposite pair are equal to n : $d((\alpha, \alpha'), v) + d((\alpha', \alpha), v) = n$.

As an example, Fig. 2 presents six connected vertices of the graph.

$v_1 = (33, 39, 17)$	$v_2 = (34, 39, 17)$	$v_3 = (35, 39, 17)$
Pairs	Pairs	Pairs
(pleasant, unpleasant) 18.67	(pleasant, unpleasant) 23	(pleasant, unpleasant) 20.58
(unpleasant, pleasant) 23.33	(unpleasant, pleasant) 19	(unpleasant, pleasant) 21.42
(pleasant, neutral) 20.25	(pleasant, neutral) 20.08	(pleasant, neutral) 22.75
(neutral, pleasant) 21.75	(neutral, pleasant) 21.92	(neutral, pleasant) 19.25
(neutral, unpleasant) 24.17	(neutral, unpleasant) 17.25	(neutral, unpleasant) 21.58
(unpleasant, neutral) 17.83	(unpleasant, neutral) 24.75	(unpleasant, neutral) 20.42
$v_4 = (33, 40, 17)$	$v_5 = (34, 40, 17)$	$v_6 = (35, 40, 17)$
Pairs	Pairs	Pairs
(pleasant, unpleasant) 16.42	(pleasant, unpleasant) 18.42	(pleasant, unpleasant) 22.42
(unpleasant, pleasant) 25.58	(unpleasant, pleasant) 23.58	(unpleasant, pleasant) 19.58
(pleasant, neutral) 18.75	(pleasant, neutral) 19	(pleasant, neutral) 22.75
(neutral, pleasant) 23.25	(neutral, pleasant) 23	(neutral, pleasant) 19.25
(neutral, unpleasant) 21.42	(neutral, unpleasant) 19.58	(neutral, unpleasant) 21.17
(unpleasant, neutral) 20.58	(unpleasant, neutral) 22.42	(unpleasant, neutral) 20.83

Fig. 2. Example of distribution of pair value in a subgraph

Example of dataset

Odorants pleasantness for individuals 1 and 2:

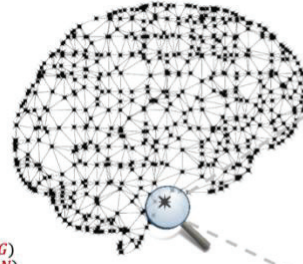
Individuals	Odorants					
	HEP	DEC	ACE	EUG	MAN	3HEX
1	pleasant	pleasant	unpleasant	unpleasant	pleasant	neutral
2	Neutral	neutral	pleasant	unpleasant	unpleasant	pleasant

Hemodynamic order of odors for individuals 1 and 2 and voxel v :

$X(v, 1, 3HEX) < X(v, 1, HEP) < X(v, 1, ACE) < X(v, 1, DEC) < X(v, 1, MAN) < X(v, 1, EUG)$
 $X(v, 2, ACE) < X(v, 2, 3HEX) < X(v, 2, EUG) < X(v, 2, DEC) < X(v, 2, HEP) < X(v, 2, MAN)$

with in green the odorants o such as $\pi_i[o] = \text{« pleasant »}$, in orange o such as $\pi_i[o] = \text{« neutral »}$ and in red o such as $\pi_i[o] = \text{« unpleasant »}$.

Attributes of the voxel v in the graph G



v	
Pairs	Values
(neutral, unpleasant)	$2/2+2/4=1.5$
(unpleasant, neutral)	$0/2+2/4=0.5$
(neutral, pleasant)	$3/3+0/4=1$
(pleasant, neutral)	$0/3+4/4=1$
(pleasant, unpleasant)	$4/6+4/4=1.66$
(unpleasant, pleasant)	$2/6+0/4=0.33$

Fig. 3. Example with 2 individuals and one voxel

To better understand the calculation of D values, we will take an example, illustrated in Fig. 3, with a toy dataset of two individuals $I = \{1, 2\}$ and a voxel v . You will find below the details of the calculations for the associated value to the pair $(\alpha, \alpha') = (\text{neutral}, \text{unpleasant})$.

First, let's count the number of odors for "neutral" and "unpleasant":

- For the individual $i = 1$:

$$|\pi_1[o] = \alpha| = |\{3HEX'\}| = 1$$

$$|\pi_1[o] = \alpha'| = |\{ACE', EUG'\}| = 2$$

- For the individual $i = 2$:

$$|\pi_2[o] = \alpha| = |\{HEP', DEC'\}| = 2$$

$$|\pi_2[o] = \alpha'| = |\{EUG, MAN'\}| = 2$$

Now, let's list the pairs (o_1, o_2) such that $\pi_i[o_1] = \text{« neutral »}$, $\pi_i[o_2] = \text{« unpleasant »}$ and $X(v, i, o_1) < X(v, i, o_2)$:

- For the individual $i = 1$:

$$|(o_1, o_2) | \pi_1[o_1] = \alpha, \pi_1[o_2] = \alpha', X(v, 1, o_1) < X(v, 1, o_2) |$$

$$= |(3HEX, ACE), (3HEX, EUG)| = 2$$

- For the individual $i = 2$:

$$|(o_1, o_2) | \pi_2[o_1] = \alpha, \pi_2[o_2] = \alpha', X(v, 2, o_1) < X(v, 2, o_2) |$$

$$= |(DEC, MAN), (HEP, MAN)| = 2$$

The value d associated to the pair $(\alpha, \alpha') = (\text{neutral}, \text{unpleasant})$ is therefore:

$$d((\alpha, \alpha'), v) = \frac{|(o_1, o_2) | \pi_1[o_1] = \alpha, \pi_1[o_2] = \alpha', X(v, 1, o_1) < X(v, 1, o_2) |}{|\pi_1[o] = \alpha| \cdot |\pi_1[o] = \alpha'|} + \frac{|(o_1, o_2) | \pi_2[o_1] = \alpha, \pi_2[o_2] = \alpha', X(v, 2, o_1) < X(v, 2, o_2) |}{|\pi_2[o] = \alpha| \cdot |\pi_2[o] = \alpha'|}$$

$$d((\alpha, \alpha'), v) = \frac{2}{1 \times 2} + \frac{2}{2 \times 2} = 1,5$$

- ii. Sub-group discovery: Exceptional attributed subgraph mining

Here, we aim to extract exceptional subgraphs using C-energetics algorithm. We are looking for sets of connected voxels attributed by pairs whose associated value is exceptionally high compared to the rest of the graph. A high value associated with a pair means that this pair is frequent for many individuals. For example, in Fig. 2, vertices v_1 , v_4 and v_5 have high pairs (neutral, pleasant) and (unpleasant, pleasant) compared to

the rest of the graph. To find these exceptional subgraphs, we use the *C-energetics* algorithm (19).

This algorithm takes as input a graph with a set of vertices V , a set of edges E , pairwise C and values D . From this graph, it returns patterns each described by a set of vertices $K \subseteq V$, a set of pairs $L \subseteq C$ that maximize a quality measure named *WRAcc* (33). The measure $WRAcc(L, K)$ quantifies the relevance of the pattern.

Let us note $G[K]$, the subgraph induced by K . We will also note $sum(X, Y)$ the sum of the values associated with the set of pairs X and the set of voxels Y :

$$sum(X, Y) = \sum_{v \in Y} \sum_{p \in X} d(p, v)$$

To evaluate the exceptionality of the values of the pairs L in the subgraph $G[K]$ compared to the rest of the graph, we use the gain function:

$$gain(L, K) = \frac{sum(L, K)}{sum(L, V)} - \frac{sum(L, V)}{sum(P, V)}$$

The gain is equal to the frequency of exceptional pairs reported in the pattern minus the expected value, i.e. the frequency of these exceptional pairs throughout the graph.

For a pair to be included in the pattern, its gain for each voxel in the pattern must be positive. The valid function is used to check this condition:

$$valid(L, K) = \bigwedge_{v \in K} \bigwedge_{p \in L} gain(p, v) > 0$$

A pattern is valid if the gain is positive for all the vertices of the pattern $v \in k$ and all the exceptional pairs of the pattern $p \in L$.

The quality of the found pattern can be measured by the *WRAcc* function:

$$WRAcc(L, K) = \begin{cases} \frac{sum(K)}{sum(V)} \times gain(L, K) & \text{if } valid(L, K) \\ 0 & \text{otherwise} \end{cases}$$

The *WRAcc* is zero if the pattern is invalid, otherwise it is equal to the support of the subgraph multiplied by the gain of the pattern. A positive gain means that the observed values for the pair L in the pattern are higher in the K voxels than expected. The presence of the support allows to maximize the number of vertices in the pattern.

From a graph $G = (V, E, P, D)$ and two thresholds σ and δ , *C-energetics* returns patterns each as an exceptional subgraph (L, K) such that 1. $|K| \geq \sigma$, 2. $G[K]$ is connected, and 3. $WRAcc(L, K) \geq \delta$.

It is important to note that this algorithm uses a closure operator to avoid redundant patterns.

iii. Filtering and validation of patterns

After having applied exceptional attributed subgraph mining to our data, we now need to focus on the patterns that are of interest for the current and/or delete those which appears non-significant following bootstrapping. To this end, we first filter the patterns in order to keep only those which interest us in this study. Among all the generated patterns, we are interested in two patterns: i/ the pleasant patterns which have the description "{(unpleasant, pleasant), (neutral, pleasant)}", i.e. pleasant stronger than neutral and unpleasant, and ii/ the unpleasant patterns with the description "{(pleasant, unpleasant), (neutral, unpleasant)}", i.e. unpleasant stronger than the two other classes. Secondly, in order to be sure that these patterns of interest are not generated by chance, we apply a validation step. For each pattern, we check that we cannot find a subgraph with a similar WRAcc by randomly drawing nodes in the graph. We calculate the distribution of the WRAcc of 10,000 connected subgraphs of the same size as the subgraph of the pattern for the same exceptional characteristics. If the WRAcc of the pattern is above the confidence interval (with $\alpha=0.025$) of this distribution then the pattern is validated otherwise it is rejected from our results. An example of pattern validation is illustrated in Fig. 4.

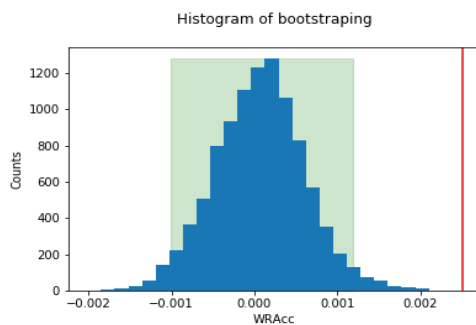


Fig. 4. Example of validation. In blue: the bootstrapping distribution. In green: the 95% percentile interval. In red: the WRAcc of the TOP1 pattern of the target "unpleasant" for the amygdala.

iv. Participation of individuals to the patterns

One of the aims of our study was to incorporate inter-individual variability into the analysis by attempting to assess how each individual contributes to the generated pattern. To do so, we calculated the involvement of each participant in the pattern using the Shapley value and also compared the involvement of individuals from different populations (depending on the odor sets used and/or depending on age) using statistical tests. In this way, we can know if a pattern concerns everyone or only a given population.

Here, to know if an individual i participates in the pattern, we took inspiration from game theory with the *Shapley value* (34). This measure is initially used to fairly distribute a payoff within a coalition. So, we used it to distribute the *gain* of the pattern between each individual in order to quantify its participation.

We are interested here in the variation (increase or decrease) of gain that the coalition can obtain with the presence of individual i . *Shapley value* corresponds to the average marginal value of this variation for all possible coalitions with individual i . The Shapley value for an individual i is denoted $\varphi_i(v)$ and is defined by:

$$\varphi_i(v) = \sum_{Z \subseteq N, i \in Z} \frac{(n - |Z|)! (|Z| - 1)!}{n!} \times (v(Z) - v(Z \setminus i))$$

where $|Z|$ is the number of individuals in the coalition Z , n the total number of individuals and $v(Z)$ is a characteristic function which gives the gain of the coalition Z . The calculation of this value is illustrated in Fig. 5 in an example of three individuals.

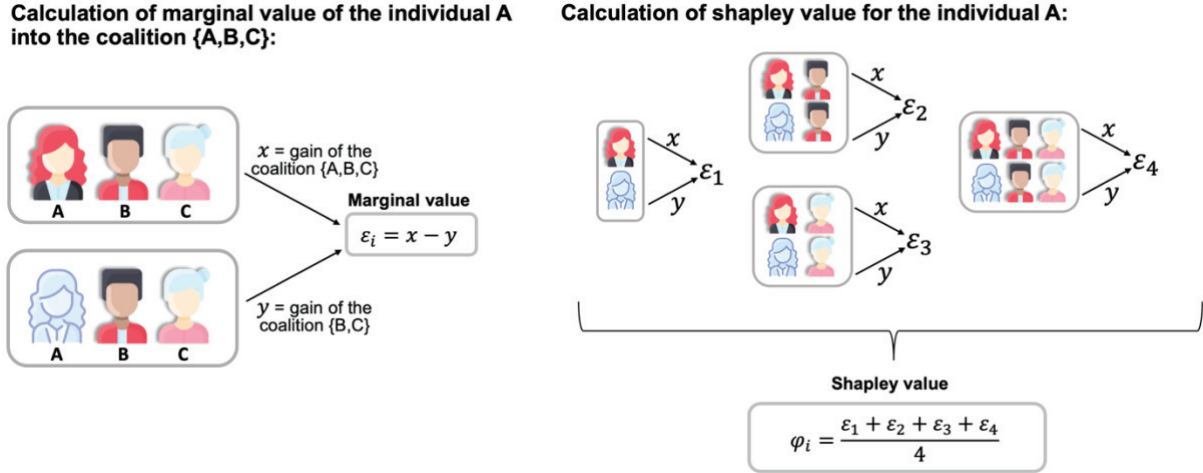


Fig. 5. Calculation of Shapley value for an individual A in a dataset with 3 individuals.

In our study, the function v is the *gain*(L, K) in the coalition-wide graph only. This graph has the same vertices K and characteristics S , the value d' associated with a pair p and a vertex v for a coalition Z is given by the function $d'(p, v, Z)$:

$$d'(p, v, Z) = \sum_{Z \subseteq N, i \in Z} \frac{|\{(o_1, o_2) \in O^2 \mid \pi_i[o_1] = \alpha, \pi_i[o_2] = \alpha', X(v, i, o_1) < X(v, i, o_2)\}|}{|\{o \in O \mid \pi_i[o] = \alpha\}| \cdot |\{o \in O \mid \pi_i[o] = \alpha'\}|},$$

with $p = (\alpha, \alpha')$

The value d' for a vertex (i.e., a voxel) v and a pair p is the sum of the average occurrence of the pair for each individual of coalition Z .

There are $2^{42} - 1$ non-empty coalitions, so calculating the shapley value is not possible. We can therefore approximate this value in polynomial time based on sampling theory with the algorithm *ApproShapley* (35). The estimate of the shapley value is the average of the marginal contributions on a sample of 15,000 random coalitions.

Once the participation of each individual is calculated for a pattern, we can then compare the participation according to other parameters such as age or gender using independent sample tests on the *shapley values* of each group. The normality of the samples is tested with the *Shapiro-Wilk* test. If the test is significant, the *Mann-Whitney* test is used. On the contrary, if the groups are normal, the equality of the variances is tested with *Levene's* test. If the variances of the two groups to be compared are not homogeneous, the *Welch t-test* is applied; otherwise, the independent *Student test* is used.

RESULTS

1. Univariate analysis

Figure 6.a depicts the functional group image of all 42 participants smelling in all odorant conditions. This analysis shows both the piriform and amygdala activities and completes the ROIs analyses in the anterior and posterior piriform cortices and the amygdala.

For the anterior portion of the piriform cortex, whereas a trend was observed for a valence effect ($F[1.84582,75.67787]=2.86457$, $p=0.06737$) (Fig. 6.b.i), the hemisphere effect was significant ($F[1,41] = 13.09049$, $p = 0.00081$) reflecting a stronger activity in the right hemisphere than in the left (Fig. 6.b.ii, 8.b.iii). Interestingly, the hemisphere-valence interaction was significant ($F[1.54964,63.53511] = 6.95765$, $p=0.00393$) reflecting increased hemodynamic activity for unpleasant odors ($m=28.37$, $IC=[21.37,35.37]$) vs. neutral odors ($m=11.31$, $IC=[3.73,18.89]$; $t(41)=-3.397$, $p=0.0015$) and pleasant odors ($m=18.95$, $IC=[9.86,28.04]$; $t(41)=2.054$, $p=0.0464$) specifically in the left hemisphere. The remaining comparisons both in the left and in the right hemisphere did not reach significance ($p>0.05$ in all cases).

For the posterior portion of the piriform cortex, whereas the valence effect was not significant ($F[1.42119,58.26893]=0.88564$, $p=0.38565$), we observed a significant hemisphere effect ($F[1,41] = 5.76482$, $p = 0.02097$) reflecting – as for the anterior piriform – a stronger activity in the right than the left hemisphere (Fig. 6.c.i, 7.c.ii). Although the hemisphere-valence interaction was significant ($F[1.54616,63.39256]=3.96834$, $p=0.03339$), mean comparisons did not reached significance neither in the left nor in the right hemisphere ($p>0.05$ in all cases).

Finally, for the amygdala, no valence ($F[1.17687,48.25180] = 1.96652$, $p = 0.16559$) (Fig. 6.d) or hemisphere effects ($F[1,41] = 2.51956$, $p = 0.12013$) were observed. Note that the hemisphere-valence interaction was also not significant ($F[1.21168,49.67871] = 1.10181$, $p = 0.31189$).

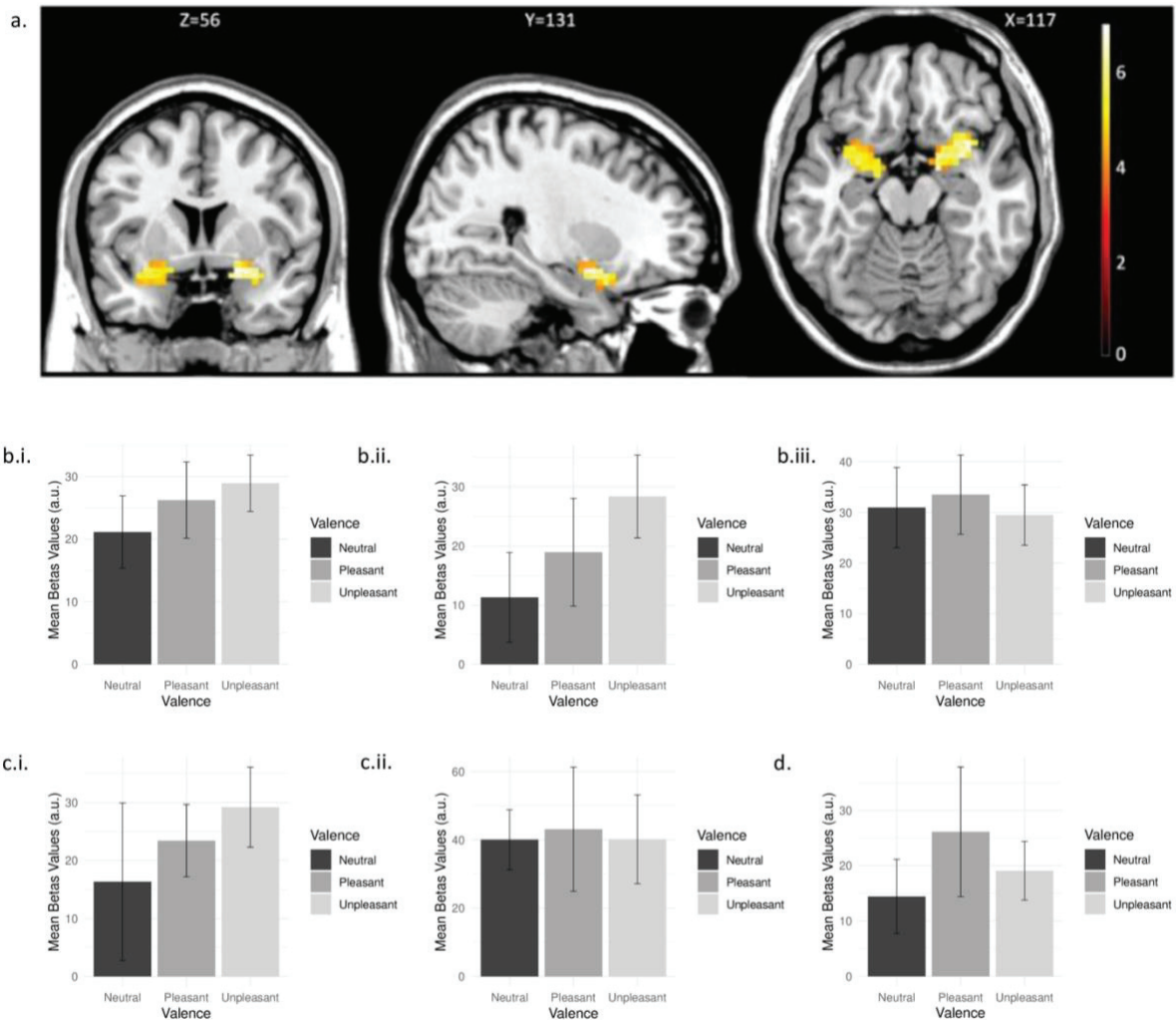


Fig. 6 - Comparison on **Ant PC**: b.i. Valence effect, b.ii. Left Hemisphere*Valence interaction, b.iii. Right Hemisphere*Valence interaction, **Post PC**: c.i. Left Hemisphere*Valence interaction, c.ii. Right Hemisphere*Valence interaction, **Amygdala**: d. Valence effect.

2. Data mining analysis

Table 1 lists all patterns extracted by the *C-energetics* algorithm ($WR_{Acc} > 0.0005$; validated by bootstrapping, see Figures S1-3). The patterns are organized by region of interest (anterior and posterior parts of the piriform cortex, amygdala), by number ordered in an ascending manner according to WR_{Acc} value. For each pattern, the number of voxels involved as well as the participation of the individuals in the pattern, the coordinates of the center of gravity and the hemisphere involved are mentioned. Finally, Table 1 also integrates the effects of the gender and age factors (p-values of the statistical tests comparing - for each pattern - the differences between women vs. men or young vs. old).

One can read from Table 1 that there is a total of 14 patterns (5 unpleasant patterns, 9 pleasant patterns) that have between 3 and 13 voxels. Figures 7, 8, and 9 illustrate these patterns for the anterior piriform cortex, the posterior piriform cortex and the amygdala respectively. The participation of the individuals in the pattern is between 42% and 67%. This participation can also be visualized in Figures 7 to 9 using Venn diagrams (generated

using *interactivenn* (12)). Furthermore, the contribution to the patterns of male/female and young/old individuals is also illustrated in the histograms of Figures 7 to 9.

The anterior piriform cortex accounted for 3 patterns, one unpleasant (in the left hemisphere, 9 voxels) and two pleasant (in the right hemisphere, 5 and 7 voxels), the unpleasant being the one with the highest *WRAcc*. There are 5 patterns in the posterior piriform cortex, two unpleasant (in the left hemisphere, 12 and 6 voxels) and three pleasant (one in the left hemisphere, 9 voxels; two in the right hemisphere, 4 and 3 voxels), the unpleasant in the left hemisphere being the one with the highest *WRAcc*. For the amygdala, we observe 6 patterns: 2 unpleasant (in the left hemisphere, 13 and 5 voxels) and 4 pleasant (two in the left hemisphere, 12 and 3 voxels; two in the right hemisphere, 8 and 7 voxels).

Finally, regardless of region of interest and pattern, no significant differences between men and women were noted. The same applies to the effect of age.

Table 1 – Generated patterns for the three ROIs.

		type	wracc	voxels	participations (%)	coordinate: mass center (X,Y,Z)	lateralization	p-value test gender	p-value test age
Anterior Piriform cortex (63 voxels)	1	unpleasant	0.01151	9	52.38	(-31.18,11.90,-17.30)	left	0.108	0.231
	2	pleasant	0.00712	7	45.24	(37.58,11.90,-17.30)	right	0.546	0.744
	3	pleasant	0.00519	5	61.90	(23.83,11.90,-20.70)	right	0.245	0.712
Posterior Piriform cortex (125 voxels)	1	unpleasant	0.00626	12	59.52	(-20.86,1.58,-13.90)	left	0.436	0.619
	2	pleasant	0.00497	9	52.38	(-31.18,5.02,-17.30)	left	0.184	0.746
	3	unpleasant	0.00329	6	66.67	(-27.74,5.02,-24.10)	left	0.336	0.221
	4	pleasant	0.00287	4	50	(27.30,1.58,-10.50)	right	0.913	0.824
	5	pleasant	0.00221	3	52.38	(16.95,5.02,-20.70)	right	0.290	0.316
Amygdala (179 voxels)	1	unpleasant	0.002499	13	54.76	(-20.86,-1.86,-13.90)	left	0.836	0.470
	2	pleasant	0.002469	12	59.52	(-27.74,-5.29,-17.30)	left	0.680	0.986
	3	pleasant	0.001826	8	50	(27.27,-1.86,-24.10)	right	0.580	0.108
	4	pleasant	0.001397	7	47.62	(23.83,-1.86,-10.50)	right	0.826	0.562
	5	unpleasant	0.001087	5	52.38	(30.71,-5.29,-17.30)	right	0.062	0.375
	6	pleasant	0.000710	3	42.86	(-24.30,1.58,-27.50)	left	0.197	0.605

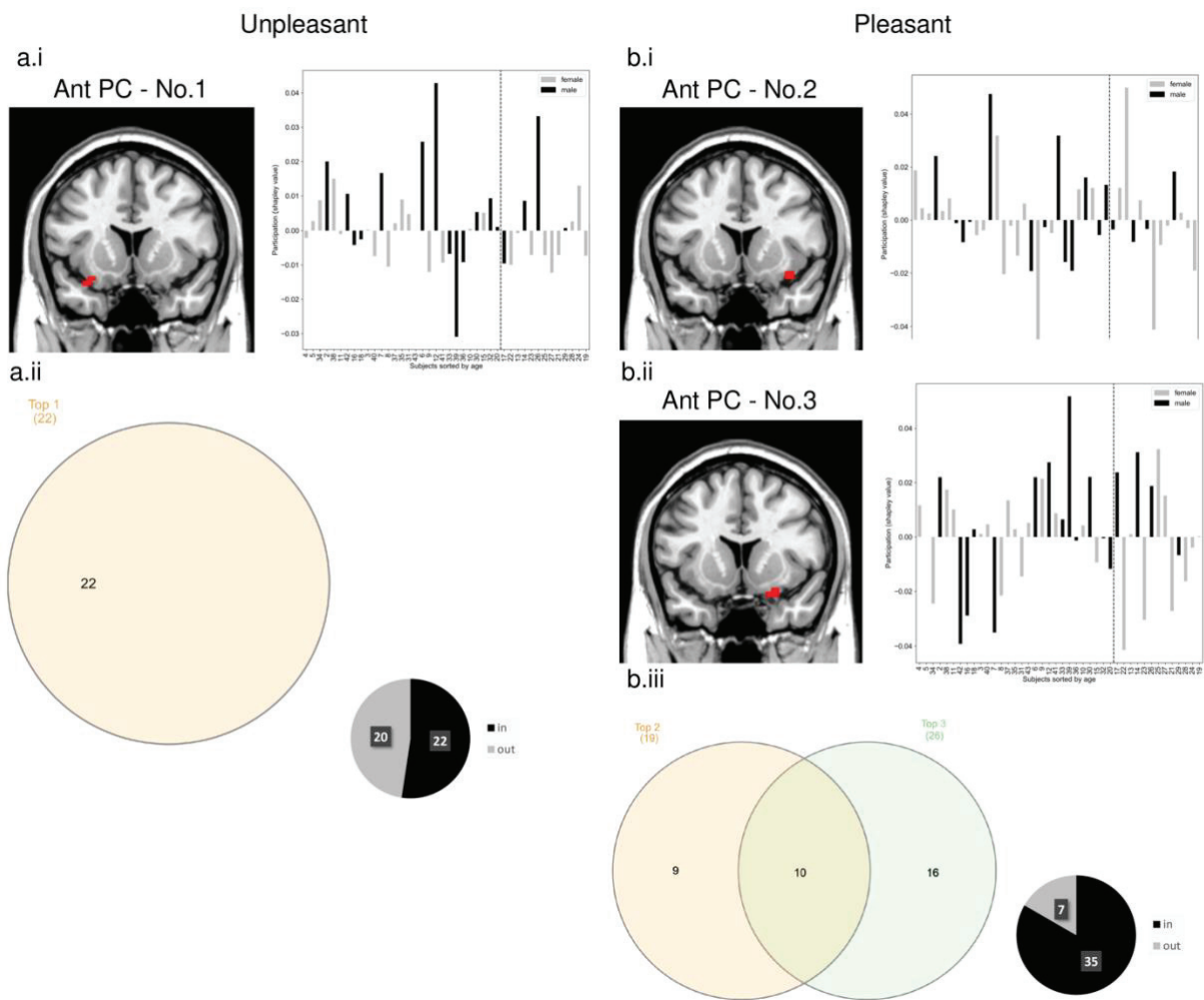


Fig. 7 – Visualization of the generated patterns for anterior piriform cortex. a) Unpleasant patterns, b) Pleasant patterns. A coronal slice illustrates each category of patterns. Moreover, a histogram shows the individuals arranged in order of age with on the left of the vertical dashed line the individuals classified as “young” and on the right the “senior” individuals. Female individuals are in grey and males in black. At the bottom right of each Venn diagram, a pie chart depicts both the number of individuals included in the Venn diagram (black) and those who are not included in these patterns (grey).

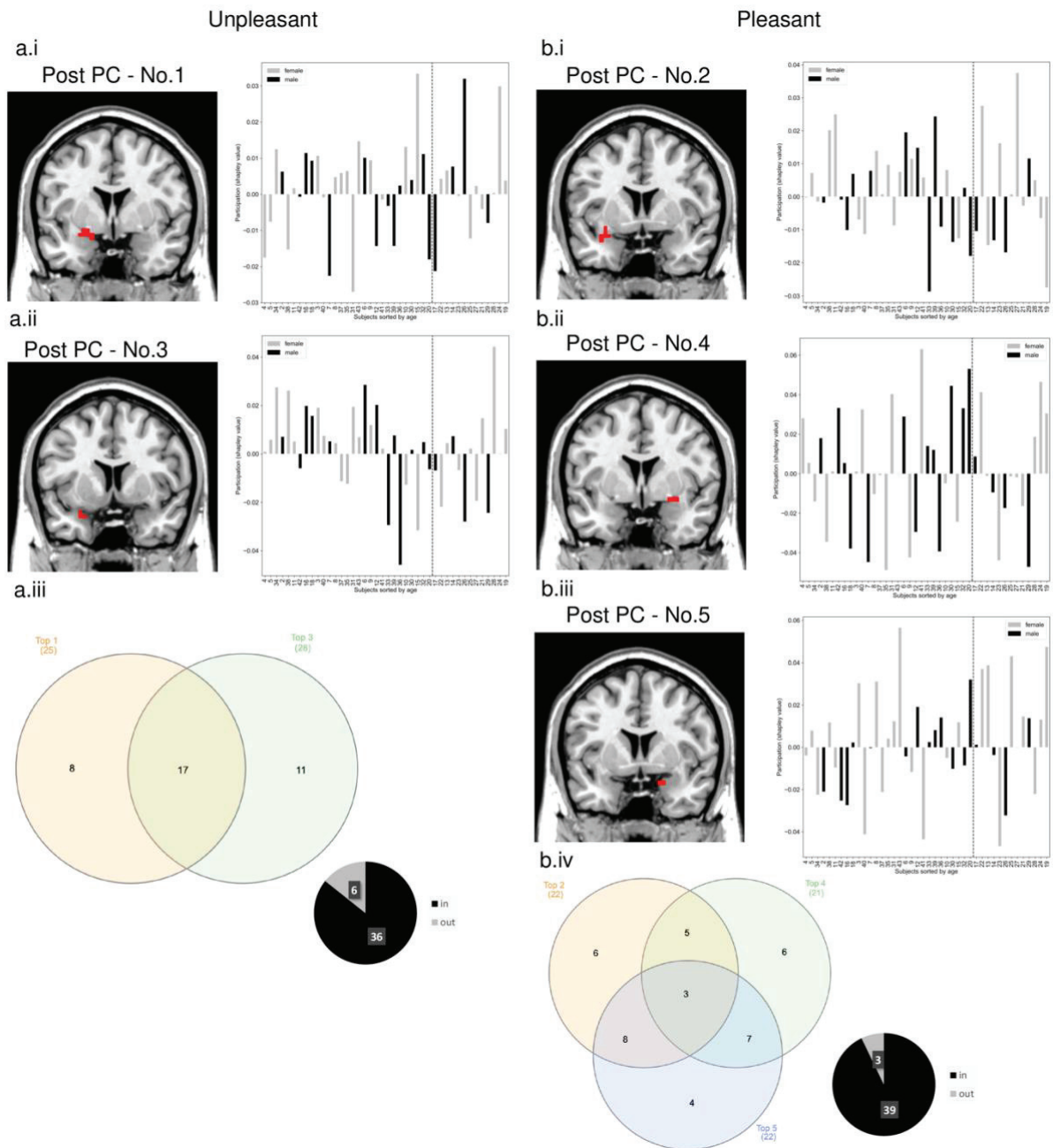


Fig. 8 - Visualization of the generated patterns for posterior piriform cortex. a) Unpleasant patterns, b) Pleasant patterns. A coronal slice illustrates each category of the patterns. Moreover, a histogram shows the individuals arranged in order of age with on the left of the vertical dashed line the individuals classified as “young” and on the right the “senior” individuals. Female individuals are in grey and males in black. At the bottom right of each Venn diagram, a pie chart depicts both the number of individuals included in the Venn diagram (black) and those who are not included in these patterns (grey).

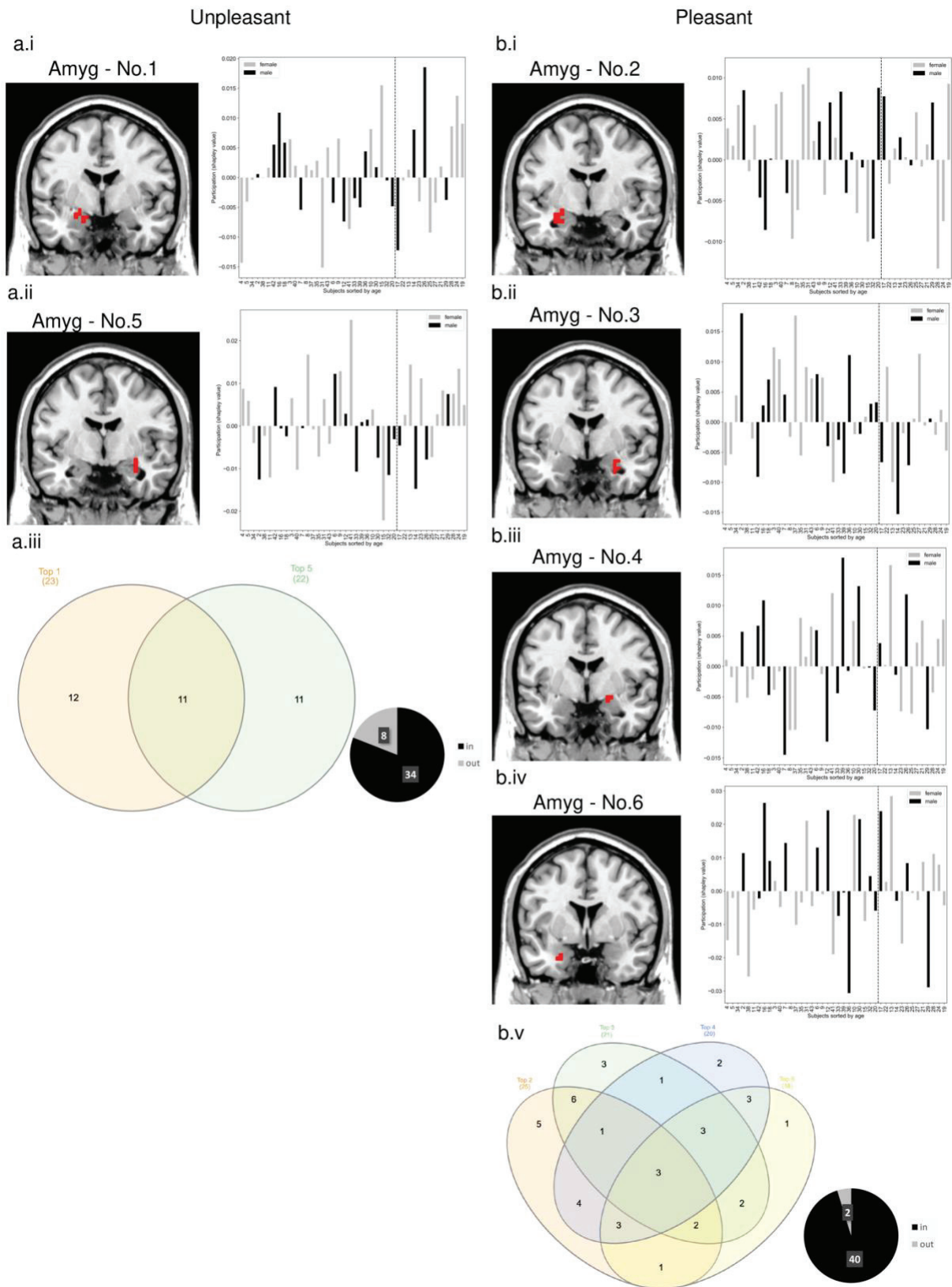


Fig. 9 - Visualization of the patterns generated for the amygdala. a) Unpleasant patterns, b) Pleasant patterns. A coronal slice illustrates each category of the patterns. Moreover, a histogram shows the individuals arranged in order of age with on the left of the vertical dashed line the individuals classified as “young” and on the right the “senior” individuals. Female individuals are in grey and males in black. At the bottom right of each Venn diagram, a pie chart depicts both the number of individuals included in the Venn diagram (black) and those who are not included in these patterns (grey).

DISCUSSION

The spatial localization of pleasant and unpleasant hedonic representations within the piriform/amygdala network is a recurrent question in recent decades (12,37–41). On the other hand, descriptive analysis methods allowing to quantify and qualify the contribution of each individual to the observed pattern - despite the inter-individual differences in perception - are still too few. Our study is at the heart of these two issues.

A univariate analysis classically used in the field revealed that unpleasant odors induced a higher hemodynamic response than neutral or pleasant odors in the anterior portion of the piriform cortex. When this result is combined with i/ psychophysical studies (using reaction times) which show that unpleasant odors are processed more rapidly than pleasant odors (42), ii/ neurophysiological studies which indicate that olfactomotor activity (sniffing) is less extensive and more rapid in response to an unpleasant odor than a pleasant odor (43,44), it can be hypothesized that the activity in the anterior piriform cortex reflects early processing of the aversive nature of odors in order to defend the organism against potentially toxic sources. This greater activity for unpleasant odors in the anterior portion of piriform cortex is in agreement with two previous studies (5,17). Interestingly, as in our study, these two publications also show that this effect is observed specifically in the left hemisphere and not the right hemisphere. The question of lateralization of olfactory processes has given rise to several works. The pioneering work of Robert Zatorre in this area showed a more pronounced involvement of the right hemisphere in olfaction (45). However, when the hedonic component is taken into account, it seems that the two hemispheres have a different contribution (8). Finally, it should be noted that the analysis of other regions of interest (posterior piriform cortex and amygdala) showed no effect of the hedonic valence on the activity of these brain areas. This result is in contradiction with some studies showing, for example, higher functional activity levels for unpleasant vs. pleasant odors in the amygdala (46) even though other works suggest that this structure could encode or represent the intensity (11) or salience (15) of the olfactory stimulus rather than its emotional valence.

The ambition of our approach was to bring additional information to this classical univariate approach via a data mining analysis that allows to develop descriptive models taking into account the contribution of each tested individual to the observed pattern. This analysis highlighted several patterns specific to unpleasant odors on the one hand, and pleasant ones on the other hand. Here, several results must be presented. First, unlike the univariate analysis, the descriptive analysis proposes patterns related to the hedonic valence in the three regions studied. Here, for the univariate analysis that was found to be significant in the anterior piriform cortex, the descriptive data mining analysis extracted a spatial localization pattern in the same structure for unpleasant odors. This pattern was localized in the left hemisphere like in the univariate analysis. In addition, the data mining approach allows a precise description of the contribution of individuals to the pattern according to various demographic factors (age, gender; here the analysis comparing men/women, young/old revealed no significant difference). Finally, in general, the number of patterns generated for pleasant odors was greater than those generated for unpleasant odors, leading to a greater percentage of individuals contributing to unpleasant patterns overall for unpleasant odors than for pleasant odors. This latter result may illustrate that there is more inter-individual agreement in judging an unpleasant odor rather than a pleasant odor. This result is consistent with previous findings from our group (25).

When one undertakes such a scientific approach consisting in linking perceptual and emotional dimensions (hedonic in our case) to a neurobiological space (fMRI), one can ask oneself how the data mining approach is positioned in relation to the classical univariate approaches. In fact, these two approaches complement each other. While the univariate approach allows for the validation of probabilistic and/or predictive models, the data mining approach is more focused on knowledge extraction by providing descriptive and especially explanatory models of the contribution of individuals to the observed pattern. In fact, because it allows a modeling that extracts from the data descriptive rules that link subgroups belonging to both hedonic and neurobiological spaces, this approach can be positioned at two levels compared to univariate analyses: 1/ upstream to generate new hypotheses in the field, scientific assumptions that will be testable with a predictive approach for example, 2/ at the same time (as in the present paper), to provide complementary information regarding spatial location and contribution of individuals to the observed pattern.

In conclusion, the main objective of the present study was to examine whether the use of exceptional pattern mining can provide new information and thus contribute to improve our understanding of the relationship between odor hedonics and central nervous system activity in the human amygdala and piriform cortex. The presented workflow allowed us to extract specific patterns of activity involved in pleasant and unpleasant odor processing and to better describe the individual variability within each of these networks. As this method can be applied to individuals who did not receive the same stimuli, it is possible to combine data from different studies but sharing a similar protocol. It is thus possible to take into account a larger number of individuals in order to better study inter-individual variability in perceptual and emotional as well as neural matter.

REFERENCES

1. Stevenson RJ. An Initial Evaluation of the Functions of Human Olfaction. *Chem Senses*. 2010 Jan 1;35(1):3–20.
2. Pinto JM. Olfaction. *Proc Am Thorac Soc*. 2011 Mar;8(1):46–52.
3. Lapid H, Shushan S, Plotkin A, Voet H, Roth Y, Hummel T, et al. Neural activity at the human olfactory epithelium reflects olfactory perception. *Nat Neurosci*. 2011 Nov;14(11):1455–61.
4. Irvani B, Schaefer M, Wilson DA, Arshamian A, Lundström JN. The human olfactory bulb processes odor valence representation and cues motor avoidance behavior. *Proc Natl Acad Sci*. 2021 Oct 19;118(42):e2101209118.
5. Bensafi M, Sobel N, Khan RM. Hedonic-Specific Activity in Piriform Cortex During Odor Imagery Mimics That During Odor Perception. *J Neurophysiol*. 2007 Dec;98(6):3254–62.
6. Zelano C, Montag J, Johnson B, Khan R, Sobel N. Dissociated Representations of Irritation and Valence in Human Primary Olfactory Cortex. *J Neurophysiol*. 2007 Mar;97(3):1969–76.
7. Jin J, Zelano C, Gottfried JA, Mohanty A. Human Amygdala Represents the Complete Spectrum of Subjective Valence. *J Neurosci*. 2015 Nov 11;35(45):15145–56.
8. Royet JP, Plailly J, Delon-Martin C, Kareken DA, Segebarth C. fMRI of emotional responses to odors: influence of hedonic valence and judgment, handedness, and gender. *NeuroImage*. 2003 Oct 1;20(2):713–28.

9. Robin O, Alaoui-Ismaïli O, Dittmar A, Vernet-Maury E. Basic Emotions Evoked by Eugenol Odor Differ According to the Dental Experience. A Neurovegetative Analysis. *Chem Senses*. 1999 Jan 1;24(3):327–35.
10. Mantel M, Ferdenzi C, Roy JM, Bensafi M. Individual Differences as a Key Factor to Uncover the Neural Underpinnings of Hedonic and Social Functions of Human Olfaction: Current Findings from PET and fMRI Studies and Future Considerations. *Brain Topogr*. 2019 Nov 1;32(6):977–86.
11. Anderson AK, Christoff K, Stappen I, Panitz D, Ghahremani DG, Glover G, et al. Dissociated neural representations of intensity and valence in human olfaction. *Nat Neurosci*. 2003 Feb;6(2):196–202.
12. Rolls ET, Kringelbach ML, De Araujo IET. Different representations of pleasant and unpleasant odours in the human brain. *Eur J Neurosci*. 2003;18(3):695–703.
13. Fulbright RK, Skudlarski P, Lacadie CM, Warrenburg S, Bowers AA, Gore JC, et al. Functional MR imaging of regional brain responses to pleasant and unpleasant odors. *Am J Neuroradiol*. 1998 Oct 1;19(9):1721–6.
14. Gottfried JA, Deichmann R, Winston JS, Dolan RJ. Functional Heterogeneity in Human Olfactory Cortex: An Event-Related Functional Magnetic Resonance Imaging Study. *J Neurosci*. 2002 Dec 15;22(24):10819–28.
15. Winston JS, Gottfried JA, Kilner JM, Dolan RJ. Integrated Neural Representations of Odor Intensity and Affective Valence in Human Amygdala. *J Neurosci*. 2005 Sep 28;25(39):8903–7.
16. Royet JP, Zald D, Versace R, Costes N, Lavenne F, Koenig O, et al. Emotional Responses to Pleasant and Unpleasant Olfactory, Visual, and Auditory Stimuli: a Positron Emission Tomography Study. *J Neurosci*. 2000 Oct 15;20(20):7752–9.
17. Gottfried JA, Dolan RJ. The Nose Smells What the Eye Sees: Crossmodal Visual Facilitation of Human Olfactory Perception. *Neuron*. 2003 Jul 17;39(2):375–86.
18. Howard JD, Plailly J, Grueschow M, Haynes JD, Gottfried JA. Odor quality coding and categorization in human posterior piriform cortex. *Nat Neurosci*. 2009 Jul;12(7):932–8.
19. Bendimerad A, Plantevit M, Robardet C. Mining exceptional closed patterns in attributed graphs. *Knowl Inf Syst*. 2018 Jul 1;56(1):1–25.
20. Fournel A, Ferdenzi C, Sezille C, Rouby C, Bensafi M. Multidimensional representation of odors in the human olfactory cortex. *Hum Brain Mapp*. 2016;37(6):2161–72.
21. Jossain P, Bessy M, Faure F, Bellil D, Landis BN, Hugentobler M, et al. Application of the European Test of Olfactory Capabilities in patients with olfactory impairment. *Eur Arch Otorhinolaryngol*. 2016 Feb 1;273(2):381–90.
22. Thomas-Danguin T, Rouby C, Sicard G, Vigouroux M, Farget V, Johanson A, et al. Development of the ETOC: A European Test of Olfactory Capabilities. :11.
23. Sezille C, Messaoudi B, Bertrand A, Jossain P, Thévenet M, Bensafi M. A portable experimental apparatus for human olfactory fMRI experiments. *J Neurosci Methods*. 2013 Aug 15;218(1):29–38.
24. Macqueen J. Some methods for classification and analysis of multivariate observations. *Proc Fifth Berkeley Symp Math Stat Probab*. 1(14):281–97.
25. Moranges M, Plantevit M, Bensafi M. Using subgroup discovery to relate odor pleasantness and intensity to peripheral nervous system reactions. *IEEE Trans Affect Comput*. 2022;1–1.

26. Esteban O, Markiewicz CJ, Goncalves M, Provins C, Kent JD, DuPre E, et al. fMRIPrep: a robust preprocessing pipeline for functional MRI [Internet]. Zenodo; 2022 [cited 2022 Aug 18]. Available from: <https://zenodo.org/record/6928849>
27. Behzadi Y, Restom K, Liu J, Liu TT. A component based noise correction method (CompCor) for BOLD and perfusion based fMRI. *NeuroImage*. 2007 Aug 1;37(1):90–101.
28. Ashburner J, Barnes G, Chen CC, Daunizeau J, Flandin G, Friston K, et al. SPM12 Manual - The FIL Methods Group (and honorary members). 2021 Oct.
29. Tzourio-Mazoyer N, Landeau B, Papathanassiou D, Crivello F, Etard O, Delcroix N, et al. Automated Anatomical Labeling of Activations in SPM Using a Macroscopic Anatomical Parcellation of the MNI MRI Single-Subject Brain. *NeuroImage*. 2002 Jan 1;15(1):273–89.
30. Rolls ET, Joliot M, Tzourio-Mazoyer N. Implementation of a new parcellation of the orbitofrontal cortex in the automated anatomical labeling atlas. *NeuroImage*. 2015 Nov 15;122:1–5.
31. Rolls ET, Huang CC, Lin CP, Feng J, Joliot M. Automated anatomical labelling atlas 3. *NeuroImage*. 2020 Feb 1;206:116189.
32. Midroit M, Chalenc̃on L, Renier N, Milton A, Thevenet M, Sacquet J, et al. Neural processing of the reward value of pleasant odorants. *Curr Biol*. 2021 Apr 26;31(8):1592-1605.e9.
33. Lavrac̃ N, Kavs̃ek B, Flach P, Todorovski L. Subgroup Discovery with CN2-SD. :36.
34. Kuhn HW, Tucker AW. Contributions to the Theory of Games. Princeton University Press; 1953. 408 p.
35. Castro J, Gómez D, Tejada J. Polynomial calculation of the Shapley value based on sampling. *Comput Oper Res*. 2009 May 1;36(5):1726–30.
36. Heberle H, Meirelles GV, da Silva FR, Telles GP, Minghim R. InteractiVenn: a web-based tool for the analysis of sets through Venn diagrams. *BMC Bioinformatics*. 2015 Dec;16(1):169.
37. Djordjevic J, Boyle JA, Jones-Gotman M. Pleasant or Unpleasant: Attentional Modulation of Odor Perception. *Chemosens Percept*. 2012 Mar 1;5(1):11–21.
38. Paradiso S, Johnson DL, Andreasen NC, O’Leary DS, Watkins GL, Boles Ponto LL, et al. Cerebral Blood Flow Changes Associated With Attribution of Emotional Valence to Pleasant, Unpleasant, and Neutral Visual Stimuli in a PET Study of Normal Subjects. *Am J Psychiatry*. 1999 Oct;156(10):1618–29.
39. Masaoka Y, Koiwa N, Homma I. Inspiratory phase-locked alpha oscillation in human olfaction: source generators estimated by a dipole tracing method. *J Physiol*. 2005;566(3):979–97.
40. Grabenhorst F, Rolls ET, Margot C, Silva MAAP da, Velazco MI. How Pleasant and Unpleasant Stimuli Combine in Different Brain Regions: Odor Mixtures. *J Neurosci*. 2007 Dec 5;27(49):13532–40.
41. Herbert C, Ethofer T, Anders S, Junghofer M, Wildgruber D, Grodd W, et al. Amygdala activation during reading of emotional adjectives—an advantage for pleasant content. *Soc Cogn Affect Neurosci*. 2009 Mar 1;4(1):35–49.
42. Bensafi M, Rouby C, Farget V, Bertrand B, Vigouroux M, Holley A. Perceptual, affective, and cognitive judgments of odors: Pleasantness and handedness effects. *Brain Cogn*. 2003 Apr 1;51(3):270–5.
43. Johnson BN, Mainland JD, Sobel N. Rapid Olfactory Processing Implicates Subcortical Control of an Olfactomotor System. *J Neurophysiol*. 2003 Aug;90(2):1084–94.
44. Bensafi M, Porter J, Pouliot S, Mainland J, Johnson B, Zelano C, et al. Olfactomotor activity during imagery mimics that during perception. *Nat Neurosci*. 2003 Nov;6(11):1142–4.

45. Zatorre RJ, Jones-Gotman M, Evans AC, Meyer E. Functional localization and lateralization of human olfactory cortex. *Nature*. 1992 Nov;360(6402):339–40.
46. Zald DH, Pardo JV. Emotion, olfaction, and the human amygdala: Amygdala activation during aversive olfactory stimulation. *Proc Natl Acad Sci*. 1997 Apr 15;94(8):4119–24.

ÉTUDE 3 : « Explorer les caractéristiques temporelles du traitement neuronal des dimensions d'éveil émotionnel et de valence hédonique dans la modalité olfactive : une approche par data mining » (en préparation)

Probing the temporal characteristics of the neural processing of arousal and valence dimensions in the olfactory modality: a data mining approach

Maëlle Moranges^{1,2}, Arnaud Fournel¹, Marc Plantevit^{2,3}, Moustafa Bensafi¹

¹Lyon Neuroscience Research Center, CNRS, INSERM, Université Lyon1

²Laboratoire d'InfoRmatique en Image et Systèmes d'information, Université Lyon1

³EPITA Research and Development Laboratory (LRDE), France

Acknowledgments: This study was granted by the ANR ChemoSim project.

Abstract

In the olfactory modality, the temporal aspects of the neural processing of the dimensions of hedonic valence and emotional arousal are still insufficiently studied. The aim of this study is to explore this issue using data mining techniques, and in particular exceptional pattern extraction techniques. For this purpose, a total of 24 volunteers participated in an EEG study in which they were asked to smell 6 odors and to describe the valence and emotional arousal felt. In line with the hemispheric lateralization theories of emotion, the preliminary analysis presented in the present paper focused on frontal electrodes located in the left (F4) and right (F3) hemispheres. The results showed earlier neural processing associated with higher frequency bands for unpleasant and exciting stimuli compared to pleasant or calming stimuli. Except for the processing of unpleasant odors, no hemispheric lateralization was observed. Finally, it should also be noted that neural responses to unpleasant odors are subject to more variability between individuals. Taken together, these data reinforce the idea of different emotional networks involved in olfactory approach and withdrawal behaviors. They also suggest that the use of a data mining approach can be of great help in developing descriptive and explanatory models of emotion and olfaction.

INTRODUCTION

The exploration of the neural bases of emotions from the electroencephalographic signal (EEG) is a subject particularly studied in the visual and auditory modalities and much less in olfaction (1–3). However, the emotional component of olfaction is an essential property for humans because it allows adopting a behavior of approach or withdrawal (4,5) towards the source of the stimulus. This adaptive reaction is triggered by evaluating whether the smell is pleasant or not, dangerous or edible. This evaluative component of the stimulus can be measured from central nervous system recordings such as EEG. An important question in the field is how affective responses to odors relate to central nervous system activity. For this purpose, Russell and Barrett's two-dimensional theory of emotion (6) can be of particular help, since emotions are described according to two basic dimensions: valence and arousal.

Several studies attempted to analyze the valence of odors from the EEG signal. Regarding the localization of the results found, odor valence seems to be represented within both the frontal and temporal lobes (7). More specifically, there is unilateral right frontal activation of brain activity during the perception of unpleasant odors (8,9). One article also reports unilateral left frontal activation for pleasant smells (10). At the temporal level, studies have shown that processing of unpleasant odors is associated with faster EEG responses than pleasant odors (11). Regarding the distinction between unpleasant and pleasant in terms of frequency bands, the frequencies δ (12), β (7–9), and γ (9,12) are recognized as being discriminating.

Some other studies set out to predict odor valence from the EEG signal. The resulting accuracy of most classification studies is better than chance but remains low (55-70%) (8,9,13). However, the authors manage to greatly improve these scores by creating a specific classifier per subject (8,9), suggesting that intra-individual information is an important aspect to consider in the analysis framework. Thus, individual variability must be taken into account and studied in order to better understand brain functioning in the perception of odor pleasantness.

The first objective of the present study is to explore the temporal aspect of the neural processing of the dimensions of hedonic valence and emotional arousal in the olfactory modality. The second objective is to include individual variability at the center of this analysis. To this end, we propose a workflow from data science to carry out a descriptive and detailed analysis of EEG activity according to arousal and valence of the odors. More precisely, we propose to use Exceptional Model Mining approaches on time-frequency data. The algorithm detects activation hotspots for different experimental conditions (e.g. pleasant, unpleasant, arousing, and calming odors). The patterns generated can then be quantified at the individual scale. This allows a finer analysis of the individual diversity very present in olfactory emotion. In line with the hemispheric lateralization theories of emotion of Davidson (14,15), the preliminary analysis presented in the present paper focused on frontal electrodes located in the left (F4) and right (F3) hemispheres.

MATERIAL AND METHODS

1. Experimental procedure

1.1. Participants

Twenty-four participants took part in the experiment (mean age 33.42 ± 9.36 years): 12 males (mean age: 33.58 ± 9.69 years) and 13 females (mean age: 33.25 ± 9.01 years). The participants were aged between 18 and 50 years old, without olfactory deficit. The absence of olfactory deficiency was confirmed using the short version of the ETOC test (16). After having read the information form, the participants provide their written consent before participating. The study was remunerated, and was in line with the Helsinki Declaration, and approved by the French CPP ethics committee.

1.2. Odorants

Six odorants provided by Sigma-Aldrich were used (CID; Odorant code; Vol/Vol dilution in mineral oil): citronellol (8842; CIT; 17.85%), heptanal (8130; HEP; 0.08%), 1-decanol (8174; DEC; 43.31%), acetophenone (7410; ACE; 0.57%), Eugenol (3314; EUG; 13.45%) and 1-propanol (1031; PRO; 0.05%).

1.3. Materials

The equipment used for the EEG recording is an Emotic EPOC + headset. This wireless headset contains 14 channels (AF3, AF4, F3, F4, F7, F8, FC5, FC6, T7, T8, P7, P8, O1 and O2) and 2 reference electrodes. We have chosen as reference electrodes, those placed behind the ear at the level of the mastoid bone. The respiratory activity related to each start of inspiration during the perception of an odor was recorded using a device allowing the synchronization with the EEG signal. This device for measuring respiratory activity is described in various publications from our group (17–19).

1.4. Experimental protocol

The 6 odors were randomly presented 15 times each. The experiment was divided into 5 sessions of 18 trials as illustrated in Fig.1.a. A trial lasted 30 seconds. Each trial began with a visual prime “Get ready to smell.”, then a screen prompted the participant to exhale and an odor was presented during the subject's inhalation until the end of the next inspiration. The subject was specifically instructed not to blink during the inspiration. The bottle containing the odorant was presented 1 cm below the nostrils by the experimenter. Five seconds after odor presentation, the participant was asked to rate the pleasantness of the odor via a response box. Next, the participant was asked to evaluate the valence and the emotional arousal dimensions of the perceived odor by placing a point in the orthonormal graph presented in Fig.1.b (6). Finally, the participant was asked to rate the intensity of the odor on a scale ranging from very low to very high (see Fig.1.c).

Note that in a preliminary session (before the 5 experimental sessions), the entire procedure was explained to the participant, and the terms “emotional arousal” and “valence” were clearly defined. Moreover, in order to be more familiar with the protocol, the participant had 3 training trials with odors (not presented during the experimental sessions). Finally, before the training and after the 5th session, the participant described their feelings for the 6 smells via a questionnaire (see Fig.1.d). In this questionnaire, the participants were asked to fill in a series of items: 6 items related to basic emotions (joy, fear, sadness, anger, disgust, surprise), 1 item related to the relaxing effect of odor and 4 items related to perceived familiarity, intensity, emotional valence and emotional arousal (data not used in the current analysis).

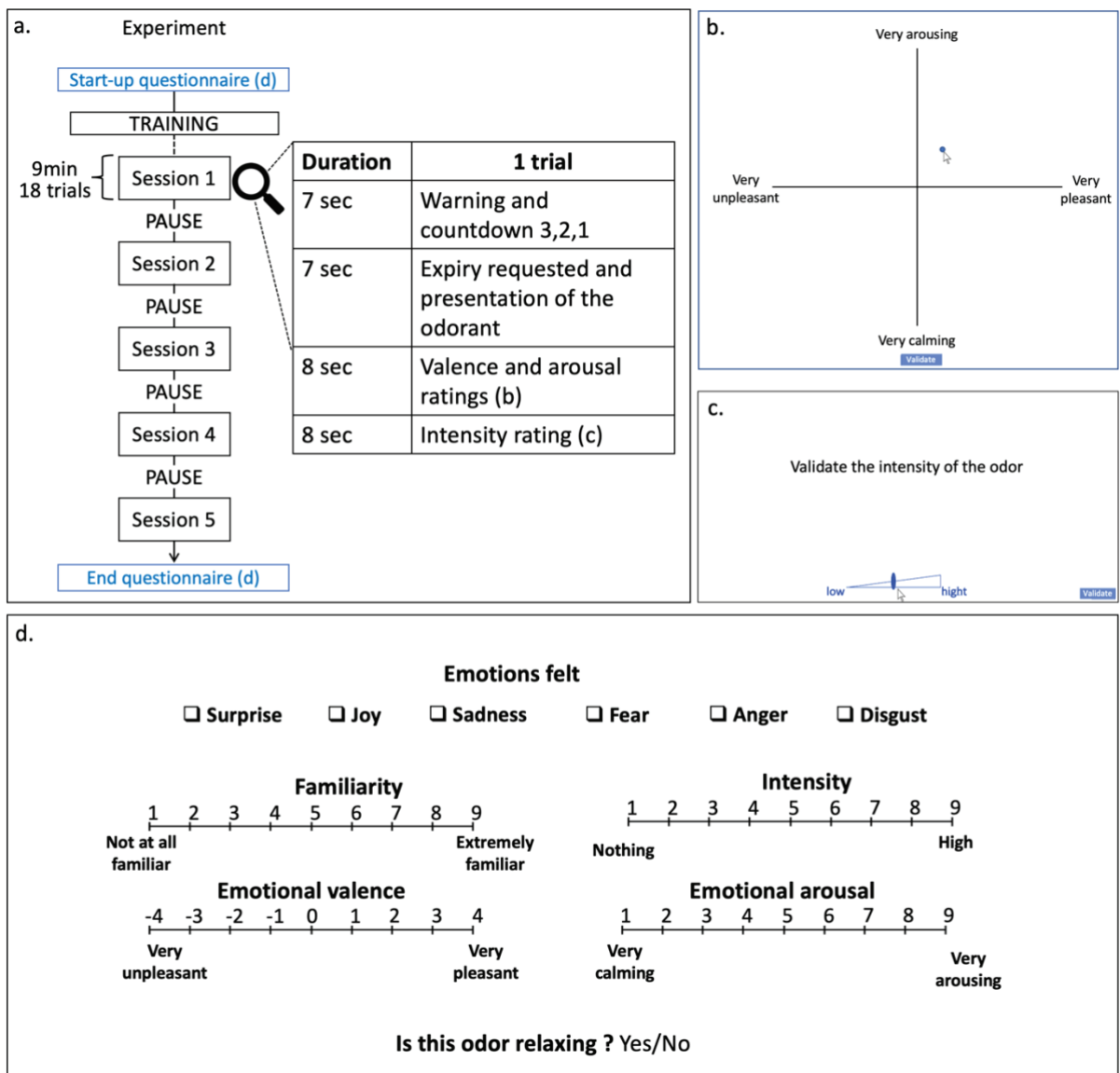


Figure 1 – a. Experimental design. b. Question screen for valence and arousal. c. Question screen for odor intensity. d. Start and end questionnaire.

2. Data analysis

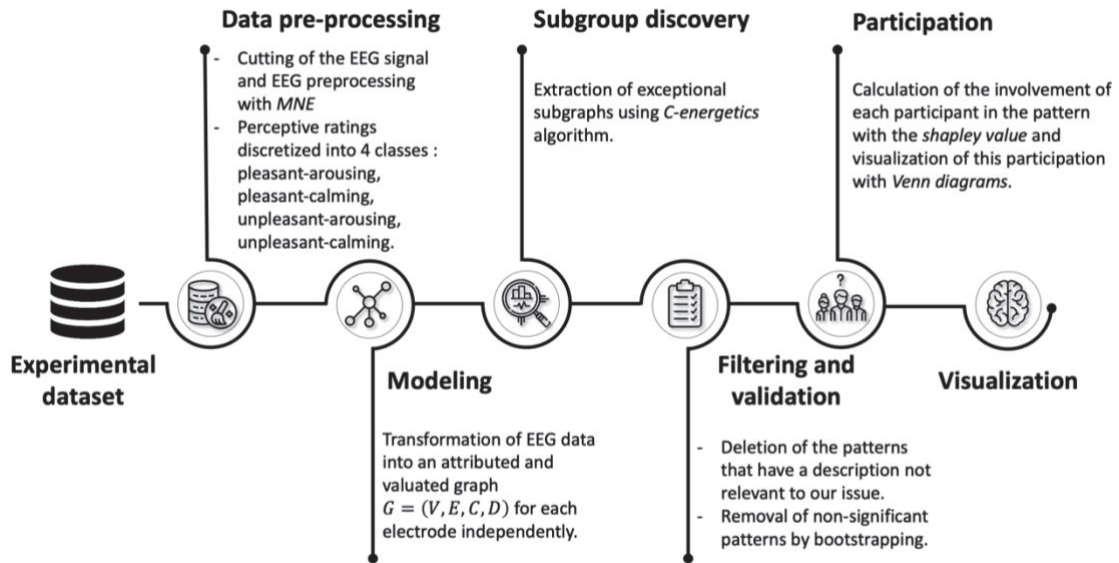


Figure 2. Steps of the workflow. The EEG raw data from the experiment was pre-processed and the subjective scores were discretized into 3 categories. Once the data was cleaned, it was transformed into an attributed and valued graph. This graph was given as input to the subgroup discovery algorithm in order to extract voxel subsets describing the different emotional states. Then, a statistical validation of the generated patterns was performed: the pattern was validated if its quality measure (*WRAcc*) was outside the confidence interval of the distribution of a bootstrap of 1,000 random draws. Finally, the participation of each individual to the pattern was calculated.

2.1. Preprocessing of EEG data

The data was preprocessed using the Python *MNE* library (20). The raw EEG data underwent a bandpass filter between 1 and 40Hz and a segmentation into epochs. Bad signal suppression and blink correction (references: AF3 and AF4) was applied. The baseline (200ms before stimuli) was removed from the signal (3 seconds after stimuli) by subtracting the mean of the reference values followed by dividing by the mean of the reference values. The average of each odor for each participant was calculated. Artifact rejection was implemented in two steps: 1/ noticeable artifacts were rejected by the experimenter, 2/ cleaned epochs were submitted to Independent Component Analysis (ICA) and each ICA component was visually inspected and removed from the source matrix before EEG signal reconstruction if identified as an artifact (eye blink, eye movement, ...). Finally, the data was analyzed using time-frequency representations (TFRs) of power for each participant and each odor within a time window starting from stimulus onset to 2.417 s post-stimulus (corresponding to the overall mean inhalation duration across all trials and participants).

2.2. Preprocessing of subjective data (odor valence and emotional arousal)

The present analysis focused on the ratings of odor valence and emotional arousal collected during each olfactory trial. For each odor, the 15 ratings given by a participant were averaged in order to obtain a pair of values (valence, arousal) for each pair (subject, odorant). Both ratings were then discretized into 4 classes as shown in Figure 3: 1/ odors with a negative valence and a high arousal categorized as “unpleasant-arousing”, 2/ odors with positive valence and high arousal categorized as “pleasant-arousing”, 3/ odors with negative valence and low arousal classified as “unpleasant-calming”, and 4/ odors with a positive valence and a low arousal classified as “pleasant-calming”.

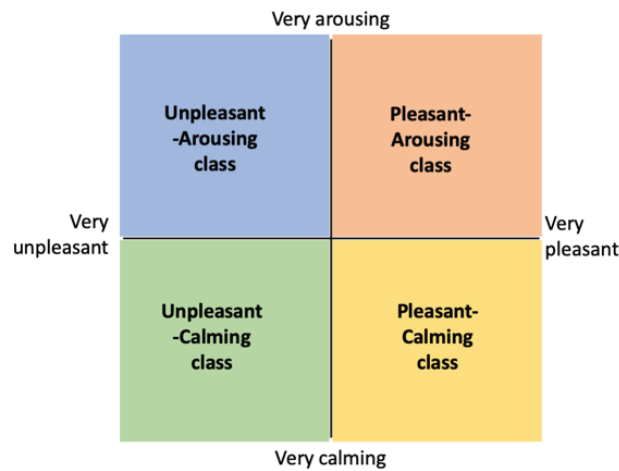


Figure 3 – Discretization of ratings of odor valence and arousal into 4 categories: “unpleasant-arousing” (upper-left), “pleasant-arousing” (upper-right), “unpleasant-calming” (lower-left) and “pleasant-calming” (lower-right).

2.3. Data modeling

The data contains a 3D table of EEG values (14 channels \times 39 frequencies \times 620 time points) associated with an emotional class for each couple (subject, odorant).

In this study, we explored each channel independently by aggregating, for each channel, the time-frequency data of each individual in the form of a graph $G_{channel} = (V, E, C, D)$.

Each time-frequency point was modeled as a vertex $v_{(t,f)}$, (t, f) being the time point pair and the associated frequency. The set of time-frequency points depicts the vertices V of the graph and the edges E connect the nearby time-frequency points. In other words, a vertex $v_{(t,f)}$ related to the time-frequency (t, f) was connected to the vertices $v_{(t-1,f)}$, $v_{(t+1,f)}$, $v_{(t,f-1)}$ and $v_{(t,f+1)}$, if these vertices exist (i.e. if $1 \leq f \leq 39$ and $1 \leq t \leq 620$).

Let $C = \{\text{pleasant-arousing}, \text{pleasant-calming}, \text{unpleasant-arousing}, \text{unpleasant-calming}\}$ be the set of the 4 categories introduced in Figure 3.

Let S be the set of the 24 subjects who took part in the experiment. We note $X(t, f, c_i, s)$, the vector of the percent change measured at time t and at frequency f for all odors of class c_i for a subject $s \in S$. The percent change is the subtraction of the mean of the baseline values followed by the division by the mean of the baseline values.

The vertices of V were described by $D = \{c_i(v_{(t,f)}) \in \mathbb{R}, \text{ with } c_i \in C \text{ and } v_{(t,f)} \in V\}$, with $c_i(v_{(t,f)}, Z) = \sum_{s \in S} \sum (X(t, f, c_i, s))$, the sum of EEG percentage change at time t and frequency f of each category c_i . The values of D can be aggregated over a set of vertices $K \subseteq V$ and a set of categories $L \subseteq C$: $sum(L, K) = \sum_{v \in K} \sum_{c_i \in L} c_i(v)$.

An example of graph with 9 vertices is illustrated in Fig. 4. The temp-frequency point represented by the vertex $v_{(1,1)}$, has neighbors: $v_{(1,2)}$ and $v_{(1,3)}$. A vector associated with the vertex gives the value of each category.

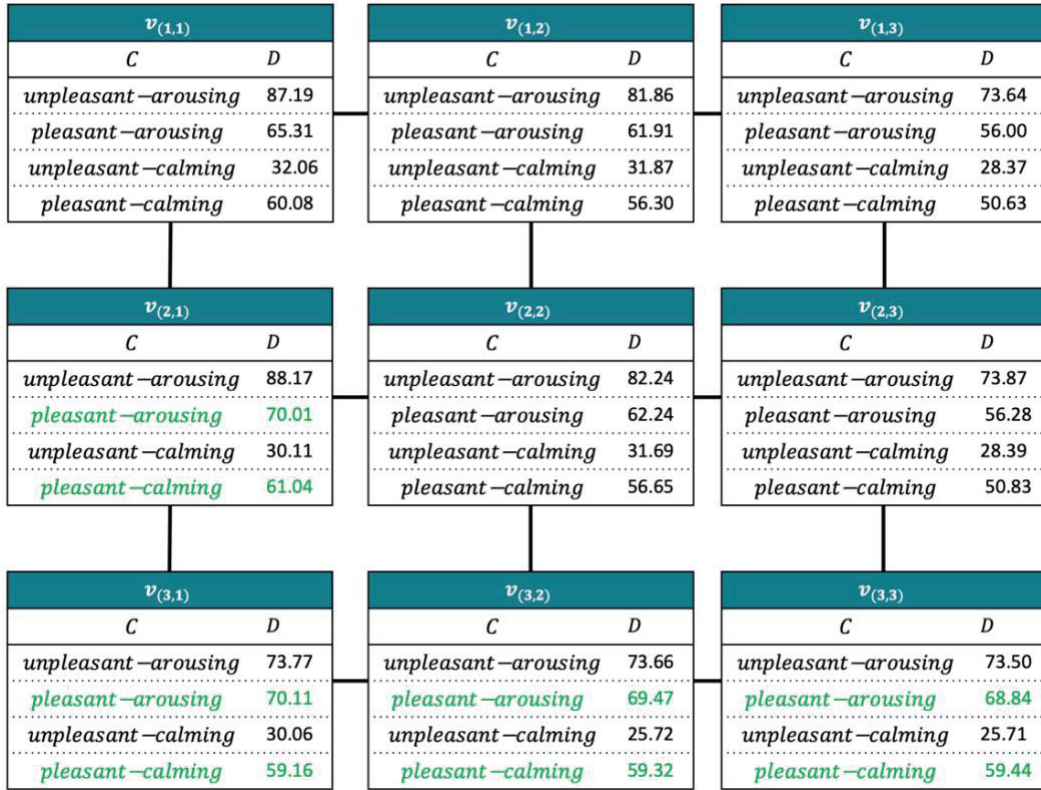


Figure 4. Example of graph G' with 9 time-frequency point. Each box illustrates a vertex. Lines connecting vertices illustrate an edge and indicate a connection.

2.4. Subgroup discovery

From a graph G , we used an algorithm named *C-energetics* algorithm (21) which extracts sets of connected vertices (e.g. time-frequency points) whose certain categories have exceptionally high values compared to the rest of the graph. If we take the example shown in Fig.4, the vertices $v_{(2,1)}$, $v_{(3,1)}$, $v_{(3,2)}$ and $v_{(3,3)}$ have high values on the categories « pleasant-arousing » and « pleasant-calming » compared to the rest of the graph. Thus, the subgraph with the vertices $K = \{v_{(2,1)}, v_{(3,1)}, v_{(3,2)}, v_{(3,3)}\}$ and the categories $L = \{\text{pleasant-arousing, pleasant-calming}\}$ is an exceptional subgraph of G .

To find these exceptional patterns *C-energetics* searches patterns that maximize a measure of quality called *WRAcc* (22). This quality measure is calculated from a gain and a validity function that we will define below.

Let us note $G[K]$, the subgraph induced by K . The exceptionality of the categories L in the subgraph $G[K]$ is evaluated by the *gain* function:

$$gain(L, K) = \frac{sum(L, K)}{sum(P, K)} - \frac{sum(L, V)}{sum(P, V)}$$

The gain is the difference between the observed value of categories and the expected value of categories. The higher the gain, the higher the exceptionality of the categories K in subgraph L .

The validity of the pattern (L, K) is given by:

$$valid(L, K) = \bigwedge_{v \in K} \bigwedge_{c_i \in L} gain(p, v) > 0$$

$valid(L, K)$ means that each vertex $v \in K$ has a positive gain for each category $c_i \in L$. $Valid(L, K)$ enforces all the vertices of the subgraph K to have a positive gain for all categories in L .

The quality of the found pattern can be measured by the *WRAcc* function:

$$WRAcc(L, K) = \begin{cases} \frac{sum(K)}{sum(V)} \times gain(L, K) & \text{if } valid(L, K) \\ 0 & \text{otherwise} \end{cases}$$

The *WRAcc* is null if the pattern is invalid, otherwise it is equal to the support of the subgraph multiplied by the gain of the pattern. A positive gain means that the observed values for the categories L in the pattern are higher in the K voxels than expected. The multiplication of this value by the support allows to maximize the number of vertices in the pattern.

From a graph $G = (V, E, P, D)$ and two thresholds σ and δ , *C-energetics* returns patterns each as an exceptional subgraph (L, K) such that 1. $|K| \geq \sigma$, 2. $G[K]$ is connected, and 3. $WRAcc(L, K) \geq \delta$.

2.5. Filtering and validation

Among all the generated patterns, we were interested in 4 types of patterns: 1) the pleasant patterns which have the categories description $K = \{\text{"pleasant-arousing"}, \text{"pleasant-calming"}\}$, 2) the unpleasant patterns which have the categories description $K = \{\text{"unpleasant-arousing"}, \text{"unpleasant-calming"}\}$, 3) the arousing patterns which have the categories description $K = \{\text{"unpleasant-arousing"}, \text{"pleasant-arousing"}\}$ and 4) the calming patterns which have the categories description $K = \{\text{"unpleasant-calming"}, \text{"pleasant-calming"}\}$.

Among the remaining patterns, to check that they were not generated by chance, we applied a validation step by bootstrapping. We draw randomly, in the graph G , 10,000 connected sub-graphs which have the same size as the tested pattern. We calculated the distribution of the *WRAcc* of these sub-graphs for the same exceptional characteristics as those of the pattern to be validated. If the *WRAcc* of the pattern was greater than the

confidence interval (with $\alpha = 0.025$) of this distribution then the pattern was validated. All non-significant patterns were removed from the study.

2.6. Participation

To calculate the contribution of each individual to the observed patterns, we used the *Shapley value* (23) from game theory. This method makes it possible to calculate the contribution of each individual to a coalition. So, we used *Shapley Value* to distribute the *gain(L, K)* of the pattern between each individual in order to quantify their individual contribution in the pattern. The *Shapley value* for an individual i is denoted $\varphi_i(v)$ and is defined by:

$$\varphi_i(v) = \sum_{Z \subseteq N, i \in Z} \frac{(n - |Z|)! (|Z| - 1)!}{n!} \times (v(Z) - v(Z \setminus i))$$

where $|Z|$ is the number of individuals in the coalition Z , n the total number of individuals and $v(Z)$ is a characteristic function which gives the gain of the coalition Z .

For the $v()$ function, we have chosen to calculate the gain at the scale of a coalition. For this purpose, the *gain(L,K)* in the coalition scale graph only. This graph has the same vertices K and the same characteristics S , the $c_i()$ values for a coalition Z are:

$$c_i(v_{(t,f)}, Z) = \sum_{Z \subseteq N, s \in Z} (X(t, f, c_i, s))$$

The number of combinations of individuals calculated being large ($2^{24} - 1$), we used an approximation of the Shapley value (24) which allows this value in polynomial time. The estimate of *Shapley value* was calculated on the average of the marginal contributions of a sample of 200 random coalitions.

Individuals with a positive contribution (i.e. a Shapley value greater than 0) were considered as pattern participants. Using the *Interactivenn* software (25), we created the Venn diagrams and study the overlaps in terms of number of individuals between the different patterns found.

RESULTS

The patterns extracted for electrodes F3 and F4 are presented in Table 1. Each line corresponds to a pattern validated by bootstrapping. The first column indicates the descriptive label of the pattern, the 2nd column the electrodes from which the pattern was extracted, the 3rd column gives the quality of the pattern with the *WRAcc* value. The 4th column presents the number of individuals included in the model. The 5th column describes the average of time and frequency of the pattern. The 6th indicates the minimum and maximum time included in the pattern and the 7th gives the minimum and maximum frequency in the pattern. The last column of the Table gives the percentage of vertices of the graphs included in each EEG frequency sub-band: delta [1-4[Hz, theta [4-8[Hz, alpha [8-12[Hz, beta [12-30[Hz and low gamma [30-39] Hz.

All these time-frequency maps are visible in Fig. 5 and 6, for each electrode and for each type of stimuli (pleasant, unpleasant, arousing, calming). The patterns have been added above in gray-transparent. In addition, the number of participants in common for a given pattern of the same category can be viewed in the form of a Venn diagram. These diagrams thus allow to illustrate whether or not the same individuals participate in the different patterns. For example, one can see in Fig. 4.a.iii that out of the 12 individuals who participate in the "F3" pleasant pattern, 11 are also present in the "F4" pattern. A single individual is present only in the "F3" pattern " and 2 are present only in pattern "F4". The remaining 10 individuals do not participate in either pattern.

Regarding the patterns for pleasant odors, one can see that we have one pattern for each of the electrodes. These two patterns are on a similar time and frequency range. They are expressed from 1.41 sec until the end of the average expiration and are mainly located in the delta and theta waves. These two patterns are supported by the same individuals (78.6% of participants in common). Note also that these are the patterns with the highest *WRAcc*.

For unpleasant odors, one pattern was found for electrode F3 and 3 patterns were found for electrode F4. The pattern in F3, is in the beta and low-gamma frequencies and runs briefly between 1.18 and 1.63 seconds. For the pattern in F4, the patterns found are not similar: a first pattern takes place immediately at the start of inspiration, lasts 0.86 seconds and is mainly located in the alpha waves. Two brief patterns in the beta waves can be seen later between 1.57 and 2.13 seconds. These 4 patterns are not supported by the same individual.

Regarding arousing odors, one pattern was found for the F3 electrode and 3 patterns were found for the F4 electrode. The pattern in F3, is mainly found in the beta frequency and takes place from the beginning of inspiration until 2.12 seconds. The 3 patterns of the F4 electrodes, describe a similar activity as the one in F3 electrode with an activity between stimulus onset and 2.13 seconds in the beta and low-gamma frequencies. Overall, these 4 patterns seem to be supported by the same individuals. With regard to calming odors, a pattern was found for each electrode in delta, theta and alpha frequencies. The two patterns begin at the same time around 1.41 seconds, the pattern of the F3 electrodes is short and lasts about 520ms, that of the F4 electrode is twice as long and lasts until the end of the expiration. The patterns have many common participants but that of the F4 electrode is supported by more participants.

Table 1. Generated patterns for F3 and F4 electrodes.

type	channel	wracc	Participations (%)	Coordinate centroid (time, freq)	Time amplitude	Freq amplitude	Percentage for each EEG band (δ , θ , α , β , γ)
pleasant	F3	0.0096	50	(1.95,4.0)	[1.41;2.42]	[1;10]	(47,44,9,0,0)
pleasant	F4	0.0055	54.17	(2.0,4.81)	[1.42;2.42]	[1;13]	(38,45,13,3,0)
unpleasant	F3	0.0008	50	(1.41,29.05)	[1.18;1.63]	[25;34]	(0,0,0,52,48)
unpleasant	F4	0.0011	58.33	(0.41,9.27)	[0.0;0.86]	[1;15]	(5,23,48,25,0)
unpleasant	F4	0.0008	50	(1.81,18.4)	[1.57;1.98]	[11;25]	(0,0,2,98,0)
unpleasant	F4	0.0006	45.83	(2.17,26.6)	[1.91;2.37]	[19;35]	(0,0,0,77,23)
arousing	F3	0.0037	50	(0.84,25.49)	[0.0;2.12]	[8;35]	(0,0,1,87,12)
arousing	F4	0.0013	62.5	(0.31,31.37)	[0.0;0.8]	[23;39]	(0,0,0,36,64)
arousing	F4	0.0007	58.33	(0.99,27.39)	[0.7;1.27]	[23;34]	(0,0,0,83,17)
arousing	F4	0.0006	58.33	(1.65,29.21)	[1.32;2.13]	[25;36]	(0,0,0,65,35)
calming	F3	0.0022	45.83	(1.71,6.63)	[1.41;1.93]	[1;12]	(20,37,38,5,0)
calming	F4	0.0028	66.66	(1.97,6.84)	[1.42;2.42]	[1;16]	(26,30,35,9,0)

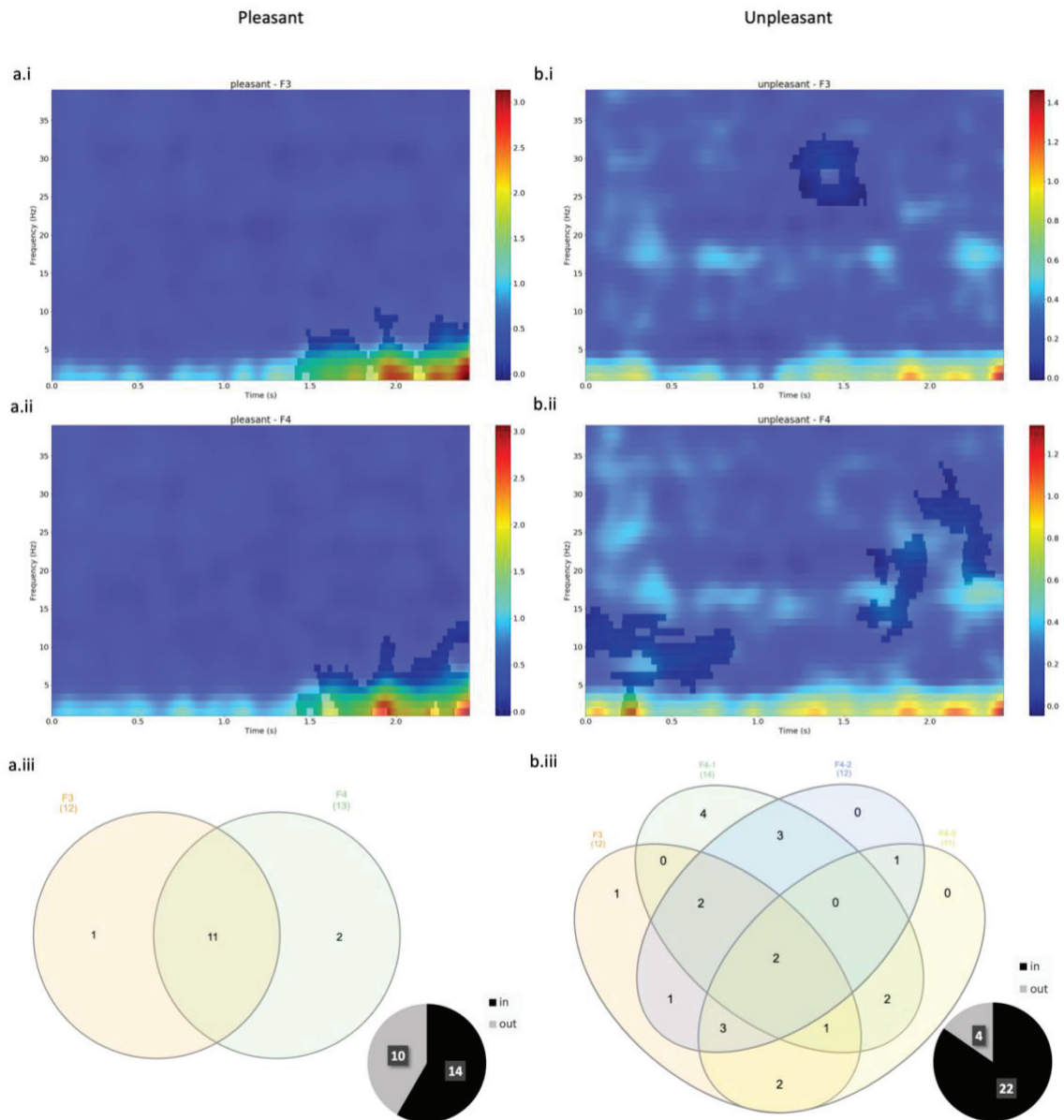


Figure 5. The patterns appear grayed out in average time-frequency maps (in percentage of change) for each for valence type (pleasant (a), unpleasant (b)). i. For F3. ii. For F4. iii. Venn diagrams of the overlap of individuals in the different patterns for each valence descriptors. At the bottom right of each diagram, a pie chart depicts both the number of individuals included in the Venn diagram (black) and those who are not included in these patterns (grey).

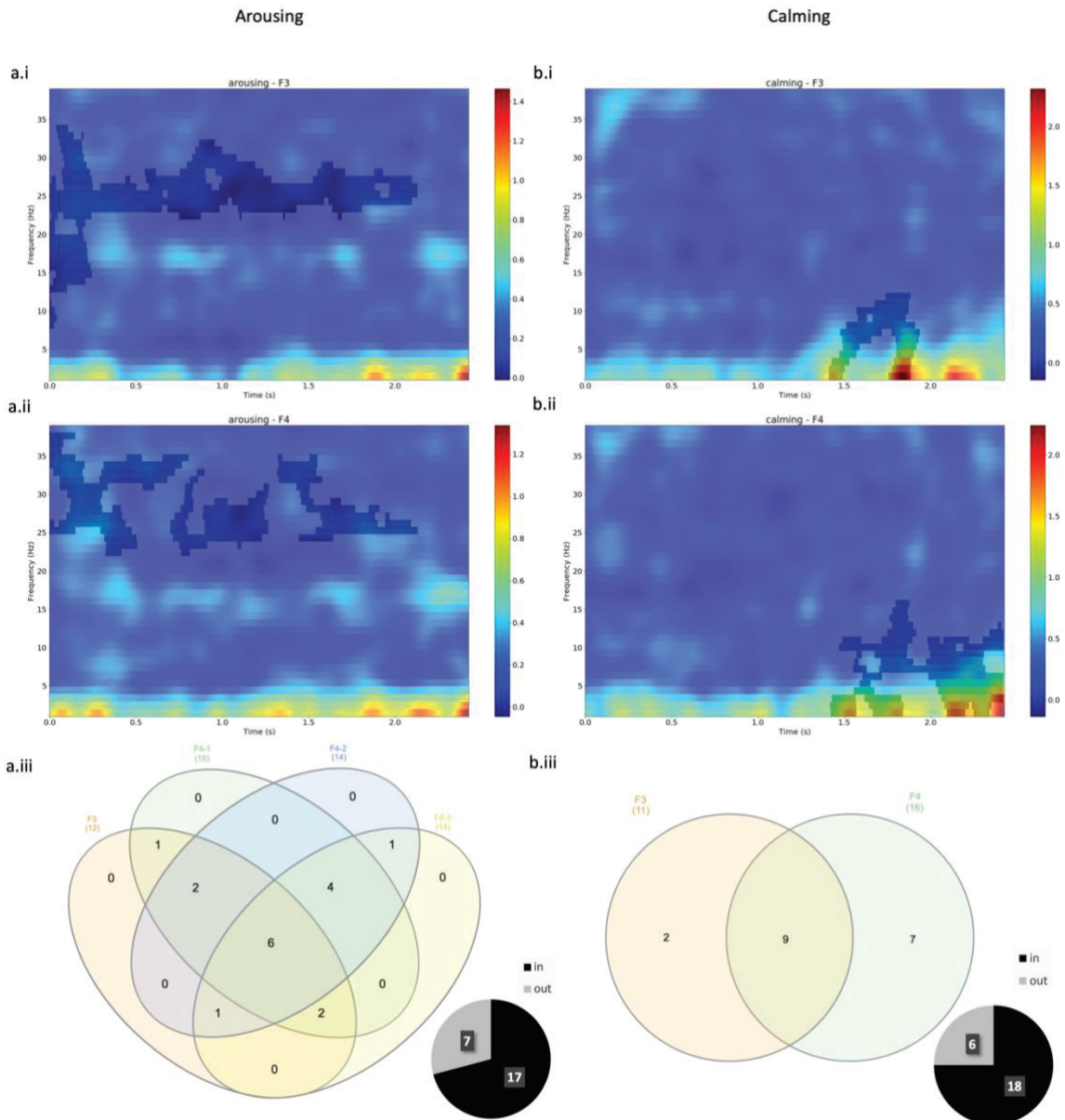


Figure 6. The patterns appear grayed out in average time-frequency maps (in percentage of change) for each for arousal category (arousing (a), calming (b)). i. For F3. ii. For F4. iii.

Venn diagrams of the overlap of individuals in the different patterns for each arousal descriptors. At the bottom right of each diagram, a pie chart depicts both the number of individuals included in the Venn diagram (black) and those who are not included in these patterns (grey).

DISCUSSION

The present study had two objectives. The first was to investigate the temporal aspect of neural processing of hedonic valence and emotional arousal dimensions in olfaction. The second was to incorporate perceptual diversity across individuals into the focus of this analysis.

A first result of interest concerns arousing odors whose perception seems to be associated with an early and bilateral neural activity. This activity is long (it lasts more than 2 seconds) and concerns the high frequencies. Conversely, calming odors are characterized by a late bilateral response in the low frequencies. The presence of beta and low-gamma waves for exciting odors could be associated with a high level of alertness of the organism. Such level of alertness may be necessary when we are confronted with such exciting stimuli. Thus, one hypothesis raised by these preliminary data could be that arousing odors would involve a rapid evaluation to determine as quickly as possible whether an approach or avoidance behavior should be adopted by the individual. This hypothesis is in agreement with a study showing a faster reaction time to ecologically relevant stimuli (edible or unpleasant) (26).

A second result of interest concerns hedonic valence. For pleasant odors, a late bilateral response is observed in the low frequencies (delta and theta rhythms) whereas the processing of unpleasant odors seems to be characterized by a hemispheric asymmetry. Indeed, for this category of odors, more patterns are found in the right hemisphere than in the left: for both hemispheres, patterns are characterized by high-frequency activities (beta and low-gamma). Nevertheless, a pattern with lower frequencies (alpha) is also present and seems to be triggered instantaneously, this pattern being the one with the largest number of individuals.

Taken together, these results are consistent with studies showing that cortical processing of unpleasant odors is faster than that of pleasant odors (11). Furthermore, an increase in high frequencies of the EEG signal in the prefrontal cortex for unpleasant odors had already been put forward in a previous study (13). Moreover, another study observed an increase in gamma waves throughout the right hemisphere for unpleasant odors (27). These results may explain why the classifiers used in previous studies distinguished pleasant from unpleasant emotions in the delta (12), beta (7–9) and gamma (9,12) frequencies: delta frequency activity could be preferentially associated with the processing of pleasant odors whereas beta and gamma frequency activity with the processing of unpleasant odors. On the other hand, although we do not find unilateral right frontal activation of brain activity for unpleasant odors as reported by different studies (8,9), we observed more patterns within the right hemisphere than in the left hemisphere which argues for the theory of a more accentuated odor processing involved in the processing of unpleasant odors in the right vs left hemisphere. Interestingly, it should be noted that among all of our participants, 12 individuals participated in the left pattern for unpleasant odors, and 21 individuals participated in the right patterns. Moreover, the right patterns are located on larger time and larger frequency amplitudes than the left pattern. An unpleasant odor potentially indicates a danger, so we can hypothesize that an unpleasant odor has an instantaneous effect and creates a state of alert in order to quickly evaluate the potential danger. Conversely, for a pleasant odor, this state of alert is not necessary and the reaction is triggered later and at lower frequencies.

In a previous study conducted by our team on individual variability in olfaction, we showed that the activity of the autonomic nervous system was specific and common for unpleasant odors, and that there were heterogeneous reactions for pleasant odors (28). We assumed in this study that unpleasant odors create a more universal response than pleasant odors. On the contrary, in the present study, we showed a common temporal reaction for pleasant odors but not for unpleasant odors. This observation leads us to raise different hypotheses.

The first is related to the analysis per se, and postulates that the more patterns we find for a given condition, the more diversity we observe. Thus, the number of patterns found would bias the results and would not really give us any indication of inter-individual variability. This first hypothesis does not seem plausible since we find as many patterns for the arousal dimension for which little diversity is observed. Conversely, we have as few calming patterns as pleasant patterns and we find a little more variability for calming odors.

The second hypothesis is that the variability found in the previous study (28), concerns a global description of the physiological activity observed in which the temporal aspect of the processing is very little represented in the data. The results of the present EEG study combined with these previous findings thus suggest that diversity can be expressed at different process levels (spatial and dimensional for psychophysiology; temporal for EEG). For pleasant odors, their processing could involve different neural networks but common temporal mechanisms between individuals. For unpleasant odors, they could involve a neural network common between individuals but whose use across individuals could be done in different ways in time. At this stage, our data cannot explain the reasons for such a temporal inter-individual difference for the perception of these unpleasant odors. It is possible that physiological or demographic factors (hormonal status, age, gender, etc.) are involved. Here, it would be interesting to study individual variability in the form of psychophysiological or fMRI time series in order to investigate whether individuals position themselves at different locations over these large temporal intervals. Our results are therefore not in contradiction with the literature but add a new temporal dimension to the process. The temporal aspect would play an important role in interindividual neurophysiological variability in the context of olfactory affect.

In conclusion, these preliminary data reinforce the idea of different emotional networks involved in olfactory approach and withdrawal behaviors. They also suggest that the use of a data mining approach can be of great help in developing descriptive and explanatory models of emotion and olfaction and their diversity. The exploration of the temporal component in the study of individual variability is a new contribution that would require further investigation in future studies. Inter-individual variability is a complex phenomenon to apprehend and depends on the nature of the data manipulated.

REFERENCES

1. Christensen L, Abdullah M. EEG emotion detection review. *EEE Conf Comput Intell Bioinforma Comput Biol CIBCB*. 2018 May 1;7.
2. Soroush MZ, Maghooli K, Setarehdan SK, Motie Nasrabadi A. A Review on EEG Signals Based Emotion Recognition. *Int Clin Neurosci J*. 2017 Oct 8;4(4):118–29.
3. Suhaimi NS, Mountstephens J, Teo J. EEG-Based Emotion Recognition: A State-of-the-Art Review of Current Trends and Opportunities. *Comput Intell Neurosci*. 2020 Sep 16;2020:e8875426.
4. Lang PJ, Reenwald MKC, Bradley MM, Hamm AO. Looking at pictures: Affective, facial, visceral, and behavioral reactions. :15.
5. Rolls ET. Taste and smell processing in the brain. In: *Handbook of Clinical Neurology* [Internet]. Elsevier; 2019 [cited 2021 Jul 13]. p. 97–118. Available from: <https://linkinghub.elsevier.com/retrieve/pii/B9780444638557000071>
6. Barrett LF, Russell JA. The Structure of Current Affect: Controversies and Emerging Consensus. *Curr Dir Psychol Sci*. 1999 Feb 1;8(1):10–4.
7. Kroupi E, Vesin J, Ebrahimi T. Subject-Independent Odor Pleasantness Classification Using Brain and Peripheral Signals. *IEEE Trans Affect Comput*. 2016 Oct;7(4):422–34.
8. Yazdani A, Kroupi E, Vesin JM, Ebrahimi T. Electroencephalogram alterations during perception of pleasant and unpleasant odors. In: *2012 Fourth International Workshop on Quality of Multimedia Experience*. 2012. p. 272–7.
9. Kroupi E, Yazdani A, Vesin JM, Ebrahimi T. EEG Correlates of Pleasant and Unpleasant Odor Perception. *ACM Trans Multimed Comput Commun Appl*. 2014 Oct 1;11(1s):13:1-13:17.
10. Kline JP, Blackhart GC, Woodward KM, Williams SR, Schwartz GER. Anterior electroencephalographic asymmetry changes in elderly women in response to a pleasant and an unpleasant odor. *Biol Psychol*. 2000 Mar 1;52(3):241–50.
11. Kato M, Okumura T, Tsubo Y, Honda J, Sugiyama M, Touhara K, et al. Spatiotemporal dynamics of odor representations in the human brain revealed by EEG decoding. *Proc Natl Acad Sci*. 2022 May 24;119(21):e2114966119.
12. Becerra MA, Londoño-Delgado E, Pelaez-Becerra SM, Serna-Guarín L, Castro-Ospina AE, Marin-Castrillón D, et al. Odor Pleasantness Classification from Electroencephalographic Signals and Emotional States. In: Serrano C. JE, Martínez-Santos JC, editors. *Advances in Computing*. Cham: Springer International Publishing; 2018. p. 128–38. (Communications in Computer and Information Science).
13. Kroupi E, Sopic D, Ebrahimi T. Non-linear EEG features for odor pleasantness recognition. In: *2014 Sixth International Workshop on Quality of Multimedia Experience (QoMEX)*. 2014. p. 147–52.
14. Tomarken AJ, Davidson RJ, Wheeler RE, Doss RC. Individual differences in anterior brain asymmetry and fundamental dimensions of emotion. *J Pers Soc Psychol*. 1992;62:676–87.
15. Sobotka SS, Davidson RJ, Senulis JA. Anterior brain electrical asymmetries in response to reward and punishment. *Electroencephalogr Clin Neurophysiol*. 1992 Oct 1;83(4):236–47.
16. Jousain P, Bessy M, Faure F, Bellil D, Landis BN, Hugentobler M, et al. Application of the European Test of Olfactory Capabilities in patients with olfactory impairment. *Eur*

Arch Otorhinolaryngol. 2016 Feb 1;273(2):381–90.

17. Ferdenzi C, Fournel A, Thévenet M, Coppin G, Bensafi M. Viewing Olfactory Affective Responses Through the Sniff Prism: Effect of Perceptual Dimensions and Age on Olfactomotor Responses to Odors. *Front Psychol* [Internet]. 2015 [cited 2022 Nov 9];6. Available from: <https://www.frontiersin.org/articles/10.3389/fpsyg.2015.01776>
18. Sezille C, Messaoudi B, Bertrand A, Jousain P, Thévenet M, Bensafi M. A portable experimental apparatus for human olfactory fMRI experiments. *J Neurosci Methods*. 2013 Aug 15;218(1):29–38.
19. Jousain P, Thevenet M, Rouby C, Bensafi M. Effect of Aging on Hedonic Appreciation of Pleasant and Unpleasant Odors. *PLOS ONE*. 2013 Apr 24;8(4):e61376.
20. Gramfort A, Luessi M, Larson E, Engemann D, Strohmeier D, Brodbeck C, et al. MEG and EEG data analysis with MNE-Python. *Front Neurosci* [Internet]. 2013 [cited 2022 Nov 2];7. Available from: <https://www.frontiersin.org/articles/10.3389/fnins.2013.00267>
21. Bendimerad A, Plantevit M, Robardet C. Mining exceptional closed patterns in attributed graphs. *Knowl Inf Syst*. 2018 Jul 1;56(1):1–25.
22. Lavrac N, Kavsek B, Flach P, Todorovski L. Subgroup Discovery with CN2-SD. :36.
23. Shapley LS, Roth AE. *The Shapley value: essays in honor of Lloyd S. Shapley*. Cambridge University Press; 1988. 342 p.
24. Kuhn HW, Tucker AW. *Contributions to the Theory of Games*. Princeton University Press; 1953. 408 p.
25. Heberle H, Meirelles GV, da Silva FR, Telles GP, Minghim R. InteractiVenn: a web-based tool for the analysis of sets through Venn diagrams. *BMC Bioinformatics*. 2015 Dec;16(1):169.
26. Boesveldt S, Frasnelli J, Gordon AR, Lundström JN. The fish is bad: Negative food odors elicit faster and more accurate reactions than other odors. *Biol Psychol*. 2010 May 1;84(2):313–7.
27. Placidi G, Avola D, Petracca A, Sgallari F, Spezialetti M. Basis for the implementation of an EEG-based single-trial binary brain computer interface through the disgust produced by remembering unpleasant odors. *Neurocomputing*. 2015 Jul 21;160:308–18.
28. Moranges M, Plantevit M, Bensafi M. Using subgroup discovery to relate odor pleasantness and intensity to peripheral nervous system reactions. *IEEE Trans Affect Comput*. 2022;1–1.

ÉTUDE 4 : « Tutoriel d'extraction de connaissances à partir de données psychophysiologiques » (accepté)

Peripheral nervous system responses to food stimuli: *Analysis using data science approaches*

Moranges M^{1,2}, Plantevit M², Bensafi M¹

¹ Lyon Neuroscience Research Center, CNRS, INSERM, Université Lyon1

² Laboratoire d'Informatique en Image et Systèmes d'information, Université Lyon1, CNRS, LIRIS UMR 5205, F-69622, Lyon, France

Abstract

In the field of food, as in other fields, the measurement of emotional responses to food and their sensory properties is a major challenge. In the present protocol, we propose a step-by-step procedure that allows a physiological description of odors, aromas and their hedonic properties. The method rooted in subgroup discovery belongs to the field of data science and especially data mining. It is still little used in the field of food and is based on a descriptive modeling of emotions on the basis of human physiological responses.

Keywords: emotions, pleasantness, odors, aromas, food stimuli, data mining, subgroup discovery

1. Introduction

The measurement of emotions is a real challenge for basic and applied research, especially in the field of food. Often, the analysis of emotions in relation to food stimuli is performed in a multimodal manner by combining subjective measures (valence or self-reported hedonic preferences) and more objective measures such as peripheral nervous system activity. Here, the challenge is to measure hedonic preferences related to odors, taste, visual aspects, texture of food products, or of the food as a whole, and to associate these verbal and declarative responses with more objective measures of emotion such as electrodermal or cardiac activity (1–5).

To date, most studies that have attempted to understand the physiological underpinnings of hedonic preferences to food sensory stimuli have used standard statistical approaches, such as comparing psychophysiological responses across different conditions (e.g., pleasant or unpleasant). Primarily used statistical methods included nonparametric tests (e.g., Wilcoxon, Kruskal Wallis) (6, 7) and parametric tests (analysis of variance, ANOVA) (8–11) depending on study design, data normality, and/or sample size.

Today, complementary approaches from data science can shed different light on these data by allowing researchers in the field to test descriptive and predictive models of the relationship between subjective preferences and psychophysiological responses. These approaches may be particularly well-suited to the format of data generated in emotion and food science as they take into account large, heterogeneous and complex data. Indeed, data science can be applied to different types of data and signals (heart rate, skin conductance, chemical data, MRI, etc.). It allows the treatment of a large number of stimuli (thousands/billion) and a large number of attributes in parallel (in our example: 4 dimensions of the skin conductance will be analyzed together).

Data science is a general term used to describe the various aspects of data processing, with the aim of extracting meaningful information and relevant knowledge (12). It includes data pre-processing (cleaning, normalization, discretization, etc.), data modeling, and data visualization. Within data modeling, we distinguish two main families of algorithms: those referring to machine learning, which are often predictive, and those belonging to data mining, which are often descriptive. Thus, whereas machine learning allows us to predict a variable (e.g. hedonic preference) on the basis of one or more variables (e.g. physiological responses) (see Note 1), data mining allows us to build descriptive models. The latter are often explanatory in the sense that they explain by explicit association rules how a pleasant food flavor is characterized physiologically compared to another less pleasant food for example.

The present protocol, inspired from a previous study (13), aims to provide researchers and students in the field of food science with a framework for using these data mining methods, especially sub-group discovery methods, in the context of research on food-related emotions. We will use electrodermal responses (e.g., skin conductance or SC) and ratings of pleasantness collected from human individuals in response to various olfactory stimuli as an example of data (14). Using these example data, we provide the user with a step-by-step

protocol that allows for a physiological description of odors and aromas characterized by their hedonic value.

2. Materials

1. A computer (*Specifications: Windows 7 or later, Mac OS X 10.11 or higher, or Linux RHEL 6/7*)
2. Python 3 SDK software: It is an interpreted, multi-paradigm and multi-platform programming language (see Note 2). It supports structured, functional and object-oriented imperative programming.
3. A Python development environment: Jupyter notebook; to be able to view the notebook created and read it as a tutorial.
4. The following libraries installed on Python:
 - a. scikit-Learn (version = 0.24.1) **(15)**: it is a library that offers various possibilities in terms data science methods
 - b. matplotlib (version = 3.1.1): to handle graphical representation
 - c. pandas (version = 1.2.3): to visualize and to manipulate tabular data
 - d. pysubgroup (version = 0.7.2) **(16)**: offers several data mining algorithms

One can install Jupyter notebook and all of the above libraries by installing a Python distribution named Anaconda and by installing the pysubgroup package with the following command : *pip install pysubgroup*.

5. A database with subjective and psychophysiological data. As an example, we will use a dataset that combines SC responses to pleasant and unpleasant odors (see Licon et al, 2018 **(14)**). The dataset is downloadable from a public repository (Dataset name: “*PsychophysioDataset.xlsx*”, available at <https://github.com/mmaelle/Psychophysio-Analysis>). Each dataset row is an observation for a specific odor with its subjective pleasantness (rated using a scale from 1, very unpleasant to 9, very pleasant) and its associated SC response described by four parameters (amplitude, latency, rise time and number of events in a period following stimulus onset). The database contains 2398 observations (109 observations for 22 individuals).

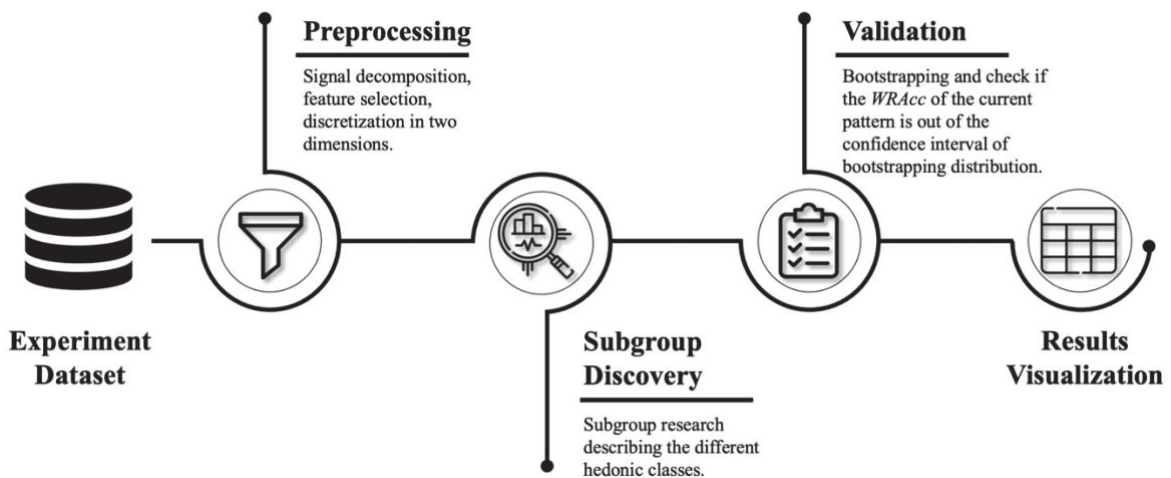


Figure 1. Steps of the workflow.

3. Methods

The protocol contains 2 steps developed below: 1/ Preprocessing, and 2/ Data mining analysis. An overview of the whole analytical process is illustrated in Figure 1.

1. Preprocessing

Preprocessing involves all the operations that precede the analysis step. Preprocessing is an important step that should not be neglected: it prepares the dataset so that it is as clean (without error), as simple as possible and is adapted to the algorithms that will be used in the following analysis (modeling and data format compatible with the input parameters of the algorithm). To do this, noisy, outlier or irrelevant data must be corrected or rejected so as not to bias the study.

1.1. Skin conductance decomposition and data importation

In a first step, we start by importing the data — available in the excel file "PsychophysioDataset.xlsx" — into a python dataframe. The data set contains the following information: participant number, CID (Compound Identification number of the odorant), latency, rise time, amplitude, number of events in a time window following stimulus onset in a time window following stimulus onset, ratings of odor pleasantness. Note that if no peak has been identified, the attributes latency, amplitude and rise time are equal to 0. Moreover, the number of events is calculated by subtracting the number of events before the presentation of the odor from the number of events after the presentation of the odor (the value can therefore be negative). For more details, refer to the second session in Licon et al (14).

```
import pandas as pd
data = pd.read_excel('PsychophysioDataset.xlsx',
                    sheet_name='Psychophysio')

print(data.head())
```

	Subject	Stimulus	Latency	Rise-Time	Amplitude	Events	Pleasantness
0	1	1	2.98046875	3.01171875	0.19736884	1	5
1	1	2	1.88671875	6.46875	0.91188065	2	4
2	1	3	2.0859375	4.74609375	0.27328656	1	0
3	1	4	1.69921875	4.515625	0.61869071	-2	5
4	1	5	0.80078125	3.4375	0.08303131	0	3

Table 1. Initial dataset

The preceding command displays the first lines of the Psychophysio dataset (Table 1).

1.2. Feature selection

In a second step, irrelevant data (such as CID) that are not directly helpful to answer our question can be removed as follows.

```
df = data[['Subject', 'Latency', 'Rise-Time', 'Amplitude', 'Events',
          'Pleasantness']].copy()
```

Incomplete data, such as trials for which participants did not provide perceptual scores or for which SC data are missing, can also be rejected.

```
df = df.drop(list(df[df['Pleasantness']==0].index))
nan_cols = [i for i in df.columns if df[i].isnull().any()]
for c in nan_cols :
    df = df.drop(list(df[df[c].isnull()].index))
```

Based on your data set, you will decide which variables are of interest and which should not be included in or should be rejected from your analysis.

1.3. Discretization

In a third step, we seek to label our data as "pleasant odors" and "unpleasant odors" for pleasantness. We will therefore discretize the self-reported scores into two classes. Having two discrete classes instead of scores allows one to deal with binary responses and to clearly separate pleasant from unpleasant odors on the one hand and weak from strong odors on the other hand. We choose here to discretize the scores using the clustering algorithm called *KMeans* (17) because this algorithm allows one to get rid of the subjectivity linked to the different scoring strategies (See Note 3). To obtain 2 groups (unpleasant and pleasant), we use $K=2$. We thus obtain the following data set:

```
print(df.head())
```

	Subject	Latency	Rise-Time	Amplitude	Events	Pleasantness_class
0	1	2.98046875	3.01171875	0.19736884	1	pleasant
1	1	1.88671875	6.46875	0.91188065	2	unpleasant
2	1	2.0859375	4.74609375	0.27328656	1	pleasant
3	1	1.69921875	4.515625	0.61869071	-2	unpleasant
4	1	0.80078125	3.4375	0.08303131	0	Pleasant

Table 2. Preprocessed dataset

The preprocessed dataset presented in Table 2 is available in csv format in the file "PsychophysioPreprocessed.csv".

2. Data mining analysis (Subgroup discovery analysis)

2.1. General information

We propose to use a data mining approach based on a subgroup discovery (SD) algorithm (18). SD allows the discovery of patterns that are discriminating for a target class. Indeed, it finds population subgroups that are statistically "most interesting" from a population of individuals (or items). These subgroups are identified by conditions on the descriptive attributes. In this regard, we seek to obtain the largest possible subpopulations that have the most unusual statistical distributional characteristics.

For example, we obtain a pattern described by the rule " $Event == 0.0 \text{ AND } Rise-Time \geq 5.35 \rightarrow unpleasant$ ". The conditions on the physiological attributes form the property of interest " $Event == 0.0 \text{ AND } Rise-Time \geq 5.35$ ". This property describes the subgroup identified as having exceptional behavior: the distribution of the target (unpleasant) is high in this subgroup compared to the rest of the dataset. This means that having a rise time greater than or equal to 5.35 seconds and at the same time a constant number of skin conductance peaks is significantly more present for the "unpleasant" trials than for the other so-called "pleasant" trials. This pattern is illustrated by the dataset depicted in Figure 2. The subgroup consists of the three colored rows in the dataset that verify the conditions present in the rule. In the dataset, rows belonging to the target class form the "positive dataset" (in bold in Figure 2) and the other rows form the "negative dataset" (in italics in Figure 2). Note that in the "positive dataset", the items that are part of the subgroup are called "positive subgroup" (or "positives_sg").

Descriptive attributes				Target
Latency	Rise-Time	Amplitude	Event	Class
3,16	1,95	0,02	-1	<i>pleasant</i>
2,98	3,01	0,20	1	<i>pleasant</i>
3,72	0,97	0,01	0	<i>pleasant</i>
3,01	6,57	0,03	0	pleasant
2,82	14,57	0,28	0	unpleasant
3,04	6,58	0,12	0	unpleasant
0,80	3,44	0,08	0	unpleasant
3,58	9,03	0,09	1	unpleasant

Subgroup
" $Event == 0.0 \text{ AND } Rise-Time \geq 5.35$ "

Positive subgroup

Positive dataset

Dataset

Figure 2. Dataset example for the pattern described by the " $Event == 0.0 \text{ AND } Rise-Time \geq 5.35 \rightarrow unpleasant$ " rule.

2.2. Processing

First, we import the data set, "pandas" library for viewing and manipulating tabular data and a library for subgroup discovery analysis called "pysubgroup".

```
import pandas as pd
import pysubgroup as ps
df = pd.read_csv("PsychophysioPreprocessed.csv", sep=',')
```

Next, we define the search space in which the algorithm should search. We remove the target column 'Pleasantness_class', so that it is not considered as a feature and does not give us the rule "Pleasant implies pleasant". Setting the nbins parameter to 20 means that the algorithm must discretize the variable values into 20 classes. The greater this value, the greater the number of rules generated and the more precise the rules.

```
searchspace = ps.create_selectors(df, nbins=20,
ignore=['Pleasantness_class'])
```

We now need to specify the target class: in our case, it is the 'Pleasantness_class' column, and we will start with the "unpleasant" odorants.

```
target = ps.BinaryTarget('Pleasantness_class', 'unpleasant')
```

Then, we create a Subgroup Discovery Task to identify the five best patterns (*result_set_size*) using the Weighted Relative Accuracy (*WRAcc*) value (*qf*). This task creates rules with a maximum of four conditions (*depth*) in the description of the subgroup. A condition can be written in different ways: either by a strict equality (eg. "Event==0" for "the number of events is zero"), or by an interval (eg. "Latency:[1.84:2.11]" for "the latency is between 1.84 and 2.11 ms"), or by a minimum or maximum value (eg. "Rise-Time>=5.35" for "the rise time is greater than or equal to 5.35ms"). When several conditions are combined, they are separated by the "AND" operator, in this case the different conditions must all be true for a trial to be included in the subgroup.

```
task = ps.SubgroupDiscoveryTask(df, target, searchspace,
result_set_size=5, depth=4, qf=ps.WRAccQF())
```

Now we can extract the patterns. The algorithm, by default, is the 'BEAM' search which performs a beam search exploration. It returns a "SubgroupDiscoveryResult" type object that can be converted into a dataframe and viewed in Table 3. We can then inspect the found subgroups and their characteristics.

```
result = ps.BeamSearch().execute(task)
unpleasant = result.to_dataframe()
```

We can also export the results to a csv file.

```
unpleasant.to_csv('unpleasant_result.csv', index=False, sep=',')
```


Then we can do the same search for “pleasant” odorants.

```
target = ps.BinaryTarget('Pleasantness_class', 'pleasant')
task = ps.SubgroupDiscoveryTask(df, target, searchspace,
    result_set_size=5, depth=4, qf=ps.WRAccQF())
result = ps.BeamSearch().execute(task)
pleasant = result.to_dataframe()
pleasant.to_csv('pleasant_result.csv', index=False, sep=',')
```

2.3. Example of results

Unpleasant pattern							
Pattern rank	Quality (WRAcc)	subgroup	size_sg	size_dataset	positives_sg	positives_dataset	lift
1	0.006387	Rise-Time>=5.35	114	2278	72	1130	1.25
2	0.003314	Amplitude>=0.50	114	2278	65	1130	1.13
3	0.003282	Event==2.0	259	2278	138	1130	1.06
4	0.002875	Latency: [1.84:2.11[114	2278	64	1130	1.11
5	0.002751	Event ==0.0 AND Rise-Time>=5.35	59	2278	36	1130	1.21
Pleasant pattern							
Pattern rank	Quality (WRAcc)	subgroup	size_sg	size_dataset	positives_sg	positives_dataset	lift
1	0.005683	Rise-Time: [1.92:2.26[113	2278	69	1148	1.23
2	0.004285	Event ==-1.0	319	2278	168	1148	1.06
3	0.003271	Rise-Time: [3.03:3.41[114	2278	64	1148	1.13
4	0.003053	Latency>=3.20	115	2278	64	1148	1.12
5	0.002541	Latency: [0.09:0.43[73	2278	42	1148	1.16

Table 3. Top 5 results obtained with *pysubgroup* library (unpleasant and pleasant patterns are depicted separately).

The top 5 results obtained for the attribute pleasantness and its targets can be viewed in Table 3 in the same format as that generated by the pysubgroup library. The best model for unpleasant has the following rule: "*Rise-Time* > = 5.35 → *unpleasant*". The subgroup size is 114 (*size_sg*) which corresponds to 5% of the dataset (*relative_size_sg* or *size_sg / size_dataset*). Of these 114 trials, 72 are classified as "unpleasant" (*positives_sg*) and 42 as "pleasant" (*size_sg - positives_sg*). So we see that we do not get a 100% accurate rule (without error), but information like "if we have a *Rise-Time* > = 5.35, then the smell is more likely to be unpleasant". To know how correct this rule is, we use quality measures.

The quality of the model is calculated as the frequency of the rule in the subgroup relative to the frequency of the rule in the entire data set. The score generally used is the Weighted Relative Accuracy (*WRAcc*) value. The *WRAcc* value is the probability of having the target in the subgroup minus the probability of having the target in the data set, multiplied by the probability of the target in the data set.

$$WRAcc = size_sg/size_dataset \times (positives_sg/size_sg - positives_dataset/size_dataset)$$

The *WRacc* value cannot exceed 0.25 or be less than -0.25. The higher the absolute value, the more exceptional the description feature in the subgroup compared to the rest of the data. A 0.25 value corresponds to the extreme case where the dataset is balanced and the subgroup -- with a size of ½ *size_dataset* -- uncovers all objects from one target without uncovering the other target (i.e., true positive rate is 1 and false positive rate is 0). Another measure of quality is shown by the parameter *lift*. The *lift* is the probability of having the target (eg. *pleasant* or *unpleasant*) in the subgroup divided by the probability of having the target in the whole dataset.

$$lift = (positive_sg / size_sg) / (positives_dataset / size_dataset)$$

If *lift* is less than 1, the rule is not interesting because the target is less frequent in the subgroup than outside. If we obtain a *lift* of 3, it means that the target is 3 times more frequent in the subgroup than in the whole data set.

We have data with a positive *WRAcc* and a *lift* greater than 1, which means that trends are found. The two quality measures are illustrated in Figure 3. One way to be sure that a pattern is meaningful and not due to chance is to perform a bootstrap validation (see Note 4).

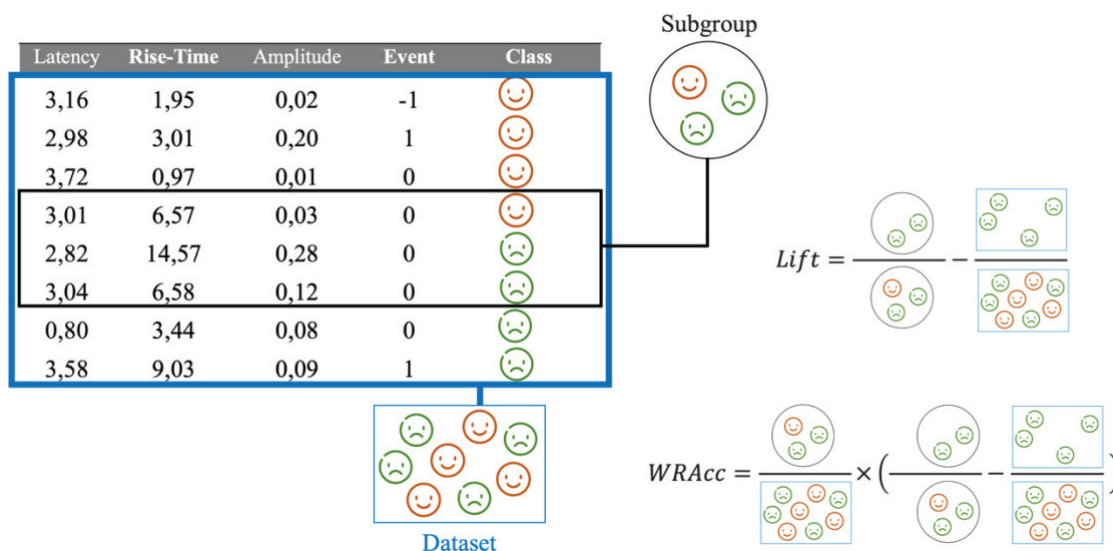


Figure 3. Illustration of quality measures

2.4. Conclusions

Subgroup discovery methods have very little implementation in olfaction and food science, but when they have (13, 19–21), they provided descriptive models linking physicochemical or physiological parameters to hedonic odor and aroma perceptions. This approach is promising and will be useful when seeking to describe perceptual phenomena and emotional reactions related to food perception from a physiological perspective.

4. Notes

Note 1. Machine learning analysis. To analyze the link between physiological and emotional responses, it is possible to use other data science methods such as predictive approaches called supervised machine learning. Supervised machine learning can be used to achieve two goals. The first is to automate a process, for example, to learn directly about a person's perceived emotions without having to ask them. The second is to know how the classifier separates the different groups in order to understand the underlying neural or physiological process better. To do this, the algorithm must be intelligible or, if it is a black box, a method must be found to allow for explanation (22). The algorithm must also have a high prediction score, which can be difficult to achieve when trying to relate complex dimensions with high variability between individuals such as physiological responses and emotional responses. To make predictions, the scikit learn library available on Python is complete and offers multiple possibilities in terms of learning methods and algorithms. In R, there are libraries corresponding to the classifier to be used (e.g. e1071, rpart, klaR, kernlab, CORElearn, Rweka, tree, caret randomForest, nnet, glmnet, gbm, rath, ipred, ROCR, mboost). In KNIME (menu Analytics/Mining), you can find a series of learning algorithms including neural network, decision tree, logistic regression. In Orange, you can find classification algorithms in the Model menu (Logistic Regression, KNN, Random Forest, SVM...).

Note 2. Other data mining tools. Besides Python, different tools and platforms exist to perform data science analysis: graphical interfaces such as Weka (23) and ELKI (24) or software such as KNIME (25) or Orange (26), with which one can compose a specific workflow by assembling one after another modules performing a specific operation. There is also an easy-to-use software with a graphical interface called Cortana (27). Cortana is also available as a plugin for KNIME. For those with good R skills, for subgroup discovery, one can use `rsubgroup` (28) on R.

Note 3. Discretization. In the literature, ratings are usually discretized by dividing the scale as used into 2 or 3 or by dividing into percentiles of equal size. However, the way in which emotions and preferences are provided is unique to each person: some people rate using a wide range of nuances, whereas other do not. It is important to separate the ratings into several categories while limiting this subjectivity. Therefore, we propose a partitioning method that is neither equi-depth nor equi-width: use *KMeans* clustering on each subject independently. This algorithm allows partitioning into *k* clusters such that the distance between intra-cluster points is minimized and the inter-cluster distance is maximized allowing the subjective data to be partitioned into categories as different as possible. We do not recommend discretizing the scores by dividing the scale into equivalent spaces (e.g. 1-5 vs. 6-10, for 2 groups and a scale ranging from 1 to 10), or partitioning into median because this method does not always reflect a person's assessment strategy. Indeed, with these methods, odors perceived in a similar way can be found in different categories and very different odors in the same category (13).

Note 4. Bootstrap validation. To filter the patterns, we can add the `min_quality` parameter to the `SubgroupDiscoveryTask()` function in order to define the minimum accepted quality. However, this is not enough to guarantee that the pattern is significant. To validate the quality of a pattern, one can make a very large number of random draws of the same size as the pattern to be validated and look at the distribution of the quality measure of all the generated draws to verify that the pattern is outside the confidence interval of that distribution. For example, by calculating the *WRAcc* of 10,000 groups for each pattern discovered, with the same support as the current pattern, drawn at random (with discount between each draw). If its *WRAcc* is outside the 90% random distribution, then the pattern is validated and can be considered as interesting. This interval validation avoids flagging subgroups indicating a *WRAcc* likely to be observed by a random subset of entities. Two examples are shown in Figure 4.

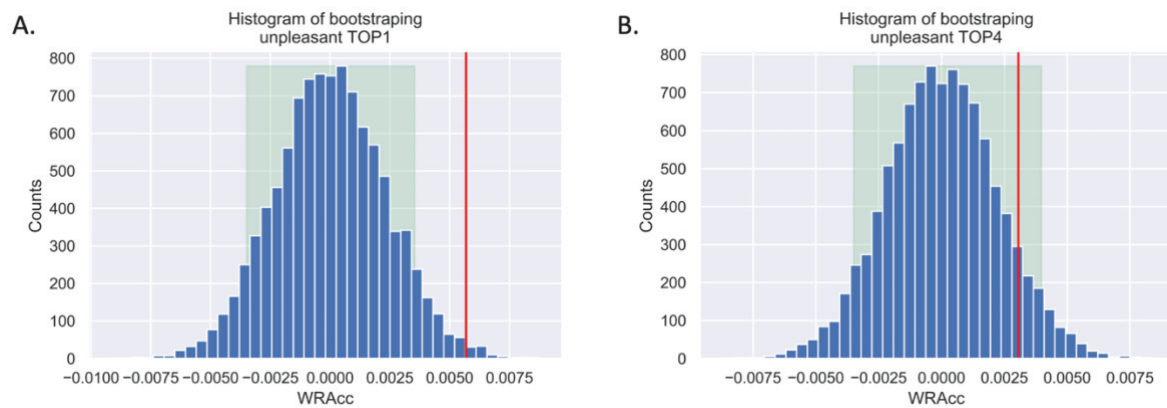


Figure 4. Examples of validation with on the left a validated pattern (A) and on the right a rejected pattern (B). The blue barplot corresponds to the distribution of the WRAcc of the random sample, the green rectangle to the confidence interval and the red vertical line to the WRAcc of the pattern to be validated.

References

1. De Wijk RA de, Kooijman V, Verhoeven RHG, et al (2012) Autonomic nervous system responses on and facial expressions to the sight, smell, and taste of liked and disliked foods. *Food Quality and Preference* 26:196–203
2. De Wijk RA, Kaneko D, Dijksterhuis GB, et al (2019) Food perception and emotion measured over time in-lab and in-home. *Food Quality and Preference* 75:170–178
3. Verastegui-Tena L, Schulte-Holierhoek A, Trijp H van, et al (2017) Beyond expectations: The responses of the autonomic nervous system to visual food cues. *Physiology & Behavior* 179:478–486
4. He W, Graaf K de Boesveld, S, and Wijk RA de (2016), Beyond liking: emotional and physiological responses to food stimuli
5. Kuoppa P, Pulkkinen K, Tarvainen MP, et al (2016) Psychophysiological responses to positive and negative food and nonfood visual stimuli. 9:78–88
6. Nardelli M, Lanata A, Francesco F di, et al (2020) Investigating complex cardiovascular dynamics during hedonic olfactory elicitation, In: 2020 11th Conference of the European Study Group on Cardiovascular Oscillations (ESGCO), pp. 1–2
7. Walsh AM, Duncan SE, Bell MA, et al (2017) Integrating implicit and explicit emotional assessment of food quality and safety concerns. *Food Quality and Preference* 56:212–224
8. Rousmans S, Robin O, Dittmar A, et al (2000) Autonomic Nervous System Responses Associated with Primary Tastes. *Chemical Senses* 25:709–718
9. He W, Boesveldt S, Graaf C de, et al (2014) Dynamics of autonomic nervous system responses and facial expressions to odors. *Front Psychol* 5
10. Muroi P, Crnjar R, and Tomassini Barbarossa I (2011) Emotional Responses to Pleasant and Unpleasant Oral Flavour Stimuli. *Chem Percept* 4:65
11. Leterme A, Brun L, Dittmar A, et al (2008) Autonomic nervous system responses to sweet taste: Evidence for habituation rather than pleasure. *Physiology & Behavior* 93:994–999
12. Moranges M, Rouby C, Plantevit M, et al (2021) Explicit and implicit measures of emotions: Data-science might help to account for data complexity and heterogeneity. *Food Quality and*

Preference 92:104181

13. Moranges M, Plantevit M, and Bensafi M (2022) Using subgroup discovery to relate odor pleasantness and intensity to peripheral nervous system reactions. 1–1
14. Licon CC, Manesse C, Dantec M, et al (2018) Pleasantness and trigeminal sensations as salient dimensions in organizing the semantic and physiological spaces of odors. 8:8444
15. Pedregosa F, Varoquaux G, Gramfort A, et al (2011) Scikit-learn: Machine Learning in Python. 12:2825–2830
16. Lemmerich F and Becker M (2019) psubgroup: Easy-to-Use Subgroup Discovery in Python, In: Brefeld, U., Curry, E., Daly, E., et al (eds.) Machine Learning and Knowledge Discovery in Databases, pp. 658–662 Springer International Publishing, Cham
17. Macqueen J Some methods for classification and analysis of multivariate observations. 1:281–297
18. Wrobel S (1997) An algorithm for multi-relational discovery of subgroups, In: Komorowski, J. and Zytkow, J. (eds.) Principles of Data Mining and Knowledge Discovery, pp. 78–87 Springer, Berlin, Heidelberg
19. Bosc G, Golebiowski J, Bensafi M, et al (2016) Local Subgroup Discovery for Eliciting and Understanding New Structure-Odor Relationships, In: Calders, T., Ceci, M., and Malerba, D. (eds.) Discovery Science, pp. 19–34 Springer International Publishing, Cham
20. Moranges M, Plantevit M, Fournel A, et al (2018) Exceptional Attributed Subgraph Mining to Understand the Olfactory Percept, In: Soldatova, L., Vanschoren, J., Papadopoulos, G., et al (eds.) Discovery Science, pp. 276–291 Springer International Publishing, Cham
21. Licon CC, Bosc G, Sabri M, et al (2019) Chemical features mining provides new descriptive structure-odor relationships. PLOS Computational Biology 15:e1006945
22. Lundberg SM and Lee S-I (2017) A Unified Approach to Interpreting Model Predictions, In: Advances in Neural Information Processing Systems, Curran Associates, Inc.
23. Frank E, Hall M, Holmes G, et al (2010) Weka-A Machine Learning Workbench for Data Mining, In: Maimon, O. and Rokach, L. (eds.) Data Mining and Knowledge Discovery Handbook, pp. 1269–1277 Springer US, Boston, MA
24. Schubert E and Zimek A (2019), ELKI: A large open-source library for data analysis - ELKI Release 0.7.5 “Heidelberg,” <http://arxiv.org/abs/1902.03616>
25. Berthold MR, Cebron N, Dill F, et al (2009) KNIME - the Konstanz information miner: version 2.0 and beyond. SIGKDD Explor Newsl 11:26–31
26. Dems̃ar J, Curk T, Erjavec A, et al Orange: Data Mining Toolbox in Python. 5
27. Meeng M and Knobbe A Flexible Enrichment with Cortana – Software Demo. 3
28. Atzmueller M and Lemmerich F (2012) VIKAMINE – Open-Source Subgroup Discovery, Pattern Mining, and Analytics, In: Flach, P.A., De Bie, T., and Cristianini, N. (eds.) Machine Learning and Knowledge Discovery in Databases, pp. 842–845 Springer, Berlin, Heidelberg

ÉTUDE 5 : « La science des données pourrait aider à tenir compte de la complexité et de l'hétérogénéité des données en sciences sensorielles et affectives » (publié)



Explicit and implicit measures of emotions: Data-science might help to account for data complexity and heterogeneity

M. Moranges^{a,b}, C. Rouby^a, M. Plantevit^{b,1}, M. Bensafi^{a,1,*}

^a Lyon Neuroscience Research Center, CNRS, INSERM, Université Lyon1, France

^b Laboratoire d'Informatique en Image et Systèmes d'information, Université Lyon1, CNRS, LIRIS UMR5205, F-69622, Lyon, France

ARTICLE INFO

Keywords:

Emotions
Physiology
Real-life situations
Data Science
Machine learning
Data mining

ABSTRACT

Measuring emotions is a real challenge for fundamental and applied research, especially in ecological contexts. de Wijk and Noldus propose combining two types of measures - explicit to characterize a specific food, and implicit -physiological- to capture the whole experience of a meal in real-life situations. This raises several challenges including development of new and miniaturized sensors and devices but also developing new ways of data analysis. We suggest a path to follow for future studies regarding data analysis: to include Data Science in the game. This field of research may enable developing predictive but also explicative models that link subjective experience of emotions and physiological responses in real-life contexts. We suggest that food scientists should go out of their comfort zone by collaborating with computer scientists and then be trained with the new tools of Data Science, which will undoubtedly enable them 1/ to better manage complex and heterogeneous data sets, 2/ to extract knowledge that will be essential to this field of research.

1. Explicit and implicit measures of emotions

Capturing emotions with an objective point of view is a challenge for research, be it fundamental or applied to food. Two main trends have emerged: quantifying subjective events such as perceptions and hedonic judgments, that require introspection and eventually training, and measuring behavioral and/or physiological cues that are less accessible to awareness. The paper of de Wijk and Noldus (this issue) recommends a combination of both kinds of measurements: mainly explicit to characterize a specific food, mainly implicit to capture the whole experience of a meal or involuntary responses to a complex food. The authors underline the specificities and difficulties that pertain to each type of measurement.

Explicit measures provide precision on many subjective facets of perception, like valence and arousal, stimulation or relaxation and allow comparisons of foods along judgments of preference, familiarity, novelty, and precise profiling of a given food. The cognitive cost of such measurements is however high, because they need introspection and reflective thinking, which capture the end of the whole perception and evaluation processes. They give very limited access about dynamics, and no access to non-conscious events that are fast and fade quickly, but may

weigh a lot in eating decisions.

On the other hand, implicit measurements follow processes that are automatic, too fast to reach awareness, that nevertheless influence decisions. Moreover, these recordings are dynamic, sensitive to both fast and basic physiological states and to slow conscious processes. They open a window on temporal processing of food perception, including the evolution of wanting and liking over time (Small et al., 2001). It is worth noting, however, that these measurements do not get direct access to the affective meaning of physiological events: they are clues, not a veridical reflection of a specific emotional processing. These clues are subject to variation between and within individuals, as attested by some discrepancies between studies. For example, whereas for some authors, the electrodermal response increases with the intensity of the stimulus (Bensafi et al., 2002b), for others it varies according to the unpleasantness (Brauchli et al., 1995) or olfactory quality (Møller & Dijksterhuis, 2003). Moreover, producing cognitive judgments like familiarity impairs some physiological responses: passive smelling and affective judgments of odors may evoke similar heart rate activity patterns, whereas familiarity judgment overwrites this autonomic response (Bensafi et al., 2002a). These differences between studies reflect a recurring phenomenon in physiological emotion studies: the great intra-

* Corresponding author at: Lyon Neuroscience Research Center, CH Le Vinatier - Bâtiment 462 - Neurocampus, 95 Boulevard Pinel, 69500 Bron, France.

E-mail address: moustafa.bensafi@cnrs.fr (M. Bensafi).

¹ co-last authors

<https://doi.org/10.1016/j.foodqual.2021.104181>

Received 17 December 2020; Received in revised form 18 December 2020; Accepted 7 January 2021

Available online 12 January 2021

0950-3293/© 2021 Elsevier Ltd. All rights reserved.

and inter-individual diversity that characterizes the data sets generated by these studies. We will come back to this notion later.

2. Combining explicit and implicit measures of emotions in real-life situations

At the end of their paper, de Wijk and Noldus raise a challenge in the field of emotional science research: to use implicit measures in real life situations as well as real consumption situations in places like restaurants, stores, cinemas and others. The authors propose different methodological and scientific locks that will need to be lifted before such ecological measures can be initiated to monitor consumer experiences 24 h a day, 7 days a week. The first lock is the miniaturization of sensors and devices to measure psychophysiological responses such as electroencephalographic (EEG) responses, cardiac and respiratory responses, electrodermal response, etc. The second lock is the development of a new type of sensors to measure psychophysiological responses. Here, the new mobile devices and portable sensors have become powerful and non-invasive enough to allow such measurements to be considered in real conditions. The third lock that the authors highlight is new software that should facilitate the analysis of subjective, behavioral and physiological data in real-world conditions. It is on this third lock that we would like to provide a path to follow for future studies.

Indeed, if we embark on such measurements of food emotions in real-life situations, we will be able to perform long-term and very large-scale experiments but - at the same time - these experiments will produce a large amount of data. This will require the use of adequate techniques to be analyzed. de Wijk and Noldus highlight the complexity of the real-life experience and there is a real need to address the subject of processing these data, which will most certainly be numerous, complex, heterogeneous and highly subject to individual variability as mentioned earlier. Actually, implicit measures are sensitive to the effects of many factors (cognitive processing as stated above but also hunger, concordance between food and packaging, etc.). In order to arrive at an in-depth model of food emotions and their physiological basis in real situations one must i) evaluate the subjective and physiological responses in their heterogeneity and complexity by considering the multiple behavioral and psychophysiological parameters involved and ii) model in an intelligible way the rules linking these physiological parameters to the consumer's subjective emotional experience. Standard statistical approaches make it possible to address the problem of the relationship between the subjective experience of food-induced emotions and their physiological foundations on the basis of a limited number of parameters and comparisons (e.g. by comparing average responses for a known food vs. a novel food). However, EEG, psychophysiological, subjective and behavioral measures remain difficult to integrate into the same analysis. Furthermore, the wide inter- and intra-individual variability that characterizes studies of emotions (food and non-food) is difficult to consider as well, especially if sample sizes exceed the standards of conventional psychophysical studies. Here, we believe that integrating Data Science in the game might help on this front.

3. What is Data Science?

If you ask this question to different actors in the field, you are likely to receive different definitions. It is actually difficult to give a consensual definition of Data Science because many scientific fields claim it as an applied subfield of their discipline. However, one can define it on the basis of its motivations: leverage existing data sources and create new ones as needed in order to extract meaningful information and actionable insights. Such motivation is not novel and falls in the scope of research in Computer Science (e.g., artificial intelligence, machine learning, data mining, database theory) or Mathematics (e.g., Statistics). Its "novelty" (or success) relies on the fact that such motivation has become ubiquitous in many scientific domains and industrial applications fostering interdisciplinary research to produce new knowledge

within a domain (Hey et al., 2009; Raghu & Schmidt, 2020), as for example food science.

Given that, such an approach borrows techniques rooted in Machine Learning, Data Mining and Statistics. Notice that the data preprocessing step should not be neglected (García et al., 2015), because it is as important as the analysis step since - if not correctly done - it can introduce important bias into the analysis and the interpretation of the results. This last point is particularly important in studies on food emotions that combine explicit measures (e.g. from subjective ratings to questionnaire data) and implicit measures (e.g. from single physiological parameters to large time-series) because the volume and the format of the data are very heterogeneous from one type of data to another, especially if the context of data acquisition is not well controlled (e.g. real-life studies).

4. How Data Science can be applied to food emotion studies?

Among the various data science analysis tasks that can be used in food emotion studies in real-life context, two important and complementary tasks can be distinguished:

1/ **Global predictive modeling** (Alpaydin, 2020), which aims to take advantage of multifaceted data to learn and construct an emotion prediction model based on different physiological measures. The ultimate goal of such a task is the automation of a process, e.g. automatically detecting whether a food evokes pleasure based on multimodal physiological responses (EEG responses, heart rate, electro-dermal response etc.) but also on factors intrinsic to the participant (e.g. satiety, knowledge about the food etc.) or extrinsic variables (e.g. context, real-life environment etc.). Such modeling emphasizes the predictive power of the model to the detriment of its interpretability (explanatory power). In other words, we will be able to predict emotion on the basis of physiology but without necessarily explaining the mechanisms underlying this relationship.

2/ **Exploratory analysis** does not aim to predict but, instead, has a strong explanatory power. This form of descriptive modeling aims to automatically discover new information about the domain (e.g. food emotions) in which the data were measured. To this end, machine discoveries are used to synergistically boost expertise of the food scientist. Here, particular attention is paid to interpretability of the results. In this domain, Subgroup Discovery (SD) (Atzmueller, 2015) (aka Exceptional Model Mining, EMM (Duivesteijn et al., 2016) makes it possible to discover subgroups (subsets) within the data that significantly deviate from the whole data (e.g., finding subgroups of individuals who have a common and particular physiological response to pleasant foods). These subgroups are identified by conditions on some feature combinations of physiological measures and are easily understandable.

One can note that making accurate predictions remains difficult in the context of human cognition, especially because this requires one to consider the intrinsic inter-individual variability of human emotions. On this point, Exploratory analysis can help to elicit new hypotheses by analyzing together both implicit and explicit measures. Especially, Subgroup discovery makes it possible to automatically discover rules between multiple implicit (physiological) measures and explicit measures for a specific subpopulation (e.g., gender = female, age <40). This type of analysis, which makes it possible to characterize subsets of individuals, can be of definite help not only to food scientists to help them better understand the diversity underlying the psychology and physiology of food emotions, but also to industrialists in the field so that their new products are more in line with the emotional expectations of specific populations of consumers. In fact, the boundary between Global predictive modeling and Exploratory analysis is not as clear-cut as we present it. New horizons are opening up in this field, in particular the field of explainable artificial intelligence (xAI) for which one expects explanation about the decision made by a black box model (e.g., deep neural networks) (Guidotti et al., 2018; Marcinkevičs & Vogt, 2020) to also support reproducibility and replicability (National Academies of

Sciences et al., 2019). Thus, developing models that predict food emotion on the basis of multiple physiological measurements obtained in real-life conditions, and especially explaining the mechanisms underlying this relationship is a realistic challenge.

5. Conclusions

In conclusion, we are in total agreement with de Wijk and Noldus to consider that combining explicit and implicit measures is a necessity if we want to better understand the food emotions evoked in real life situations. Explicit and implicit measures are complementary and will likely enable approaching access to behavioral and physiological underpinnings of emotions. The problem is whether we really benefit from this combination of data with classical statistical approaches. Here, we suggest that the food scientist must go out of his/her comfort zone by collaborating with computer scientists and then be trained in the new tools of Data Science, which will undoubtedly enable him/her 1/ to better manage complex and heterogeneous data sets, 2/ to extract knowledge that will be essential to this field of research.

Funding

This study was granted by the ANR ChemoSim project.

CRedit authorship contribution statement

M. Moranges: Conceptualization, Methodology, Writing - original draft, Writing - review & editing. **C. Rouby:** Conceptualization, Methodology, Writing - original draft, Writing - review & editing. **M. Planchevit:** Conceptualization, Methodology, Writing - original draft, Writing - review & editing. **M. Bensafi:** Conceptualization, Methodology, Writing - original draft, Writing - review & editing, Funding acquisition, Project administration.

Declaration of Competing Interest

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.

Acknowledgments

We wish to thank the ANR for their support (ANR-ChemoSim project).

References

- Alpaydin, E. (2020). Introduction to machine learning. MIT press.
- Atzmueller, M. (2015). Subgroup discovery. *Wiley Interdisciplinary Reviews: Data Mining and Knowledge Discovery*, 5(1), 35–49.
- Bensafi, M., Rouby, C., Farget, V., Bertrand, B., Vigouroux, M., & Holley, A. (2002a). Influence of affective and cognitive judgments on autonomic parameters during inhalation of pleasant and unpleasant odors in humans. *Neuroscience Letters*, 319(3), 162–166.
- Bensafi, M., Rouby, C., Farget, V., Bertrand, B., Vigouroux, M., & Holley, A. (2002b). Autonomic nervous system responses to odours: The role of pleasantness and arousal. *Chemical Senses*, 27(8), 703–709.
- Brauchli, P., Rüegg, P. B., Etzweiler, F., & Zeier, H. (1995). Electrocortical and autonomic alteration by administration of a pleasant and an unpleasant odor. *Chemical Senses*, 20(5), 505–515.
- Duivesteijn, W., Feelders, A. J., & Knobbe, A. (2016). Exceptional model mining. *Data Mining and Knowledge Discovery*, 30(1), 47–98.
- García, S., Luengo, J., & Herrera, F. (2015). Data preprocessing in data mining. Springer.
- Guidotti, R., Monreale, A., Ruggieri, S., Turini, F., Giannotti, F., & Pedreschi, D. (2018). A survey of methods for explaining black box models. *ACM Computing Surveys (CSUR)*, 51(5), 1–42.
- Hey, T., Tansley, S., Tolle, K., & others. (2009). The fourth paradigm: Data-intensive scientific discovery (Vol. 1). Microsoft research Redmond, WA.
- Marcinkevičs, R., & Vogt, J. E. (2020). Interpretability and Explainability: A Machine Learning Zoo Mini-tour. ArXiv Preprint ArXiv:2012.01805.
- Møller, P., & Dijksterhuis, G. (2003). Differential human electrodermal responses to odours. *Neuroscience Letters*, 346(3), 129–132.
- National Academies of Sciences, Engineering, Medicine, & others. (2019). Reproducibility and replicability in science. National Academies Press.
- Raghu, M., & Schmidt, E. (2020). A survey of deep learning for scientific discovery. ArXiv Preprint ArXiv:2003.11755.
- Small, D. M., Zatorre, R. J., Dagher, A., Evans, A. C., & Jones-Gotman, M. (2001). Changes in brain activity related to eating chocolate: From pleasure to aversion. *Brain*, 124(9), 1720–1733.

DISCUSSION

L'étude de la neurophysiologie des émotions dans le domaine de la perception olfactive a fait l'objet de nombreuses recherches ces dernières décennies [162,178,183,193,248,251,256,284–293]. Cependant, la diversité interindividuelle, très présente en matière d'olfaction, ainsi que ses facteurs de variations, ont jusqu'ici été très peu pris en compte dans ces études, ceci par manque d'outils méthodologiques adéquats.

L'objectif de la présente thèse était double. Dans un premier objectif, nous avons tenté d'apporter des éléments de connaissance sur les représentations des odeurs et de leur composante hédonique au sein du système nerveux. Dans un second objectif, nous avons intégré à l'approche neuroscientifique classique une approche méthodologique issue de l'intelligence artificielle et de la fouille de données afin de fournir des modèles descriptifs qui expliquent les effets des odeurs agréables et désagréables sur le fonctionnement du système nerveux central et périphérique. Ces modèles ont permis de prendre en compte la diversité interindividuelle en matière d'odeurs et d'émotions.

L'application d'algorithmes de découverte de sous-groupes à des mesures du système nerveux périphérique a donné lieu à des descriptions du caractère plaisant/déplaisant et de l'intensité d'une odeur à travers les réactions physiologiques individuelles (étude 1). Ensuite, nous avons utilisé la découverte de sous-graphes exceptionnels, pour analyser l'activation du système nerveux central à partir de données IRMf et comprendre la représentation cérébrale des odeurs plaisantes et déplaisantes au sein de deux structures d'intérêt : le cortex piriforme et l'amygdale (étude 2). Dans une troisième étude (étude 3), la découverte de sous-graphes exceptionnels a été appliquée sur un jeu de données d'électroencéphalographie (EEG) afin d'étudier l'aspect temporel de l'activité du système nerveux central. Ensuite, nous avons proposé un protocole d'extraction de connaissance à partir de données physiologiques de réponses électrodermales à des odeurs plaisantes et déplaisantes (étude 4). Pour finir, dans le cadre d'un commentaire d'article, nous avons présenté une réflexion sur la place future de la science des données dans le champ des neurosciences (étude 5).

Dans la partie qui va suivre, nous allons résumer et mettre en commun les résultats de ces différentes études. Nous allons tout d'abord poser la question de savoir en quoi les approches descriptives provenant de la science des données peuvent être utiles pour mieux comprendre le lien entre les émotions et l'olfaction. Ensuite, nous décrirons les connaissances que nous ont permis de découvrir ces approches sur la neurophysiologie de l'émotion dans le cadre de la perception olfactive. Enfin, nous aborderons les limites rencontrées durant ces travaux de thèses ainsi que les perspectives envisagées.

I. La science des données : un domaine d'analyse adéquat pour l'exploration du lien entre émotions et olfaction

Est-ce que les outils d'analyses descriptives sont utiles pour mieux comprendre la neurophysiologie de l'émotion ainsi que sa diversité dans le cadre de la perception olfactive ? Pour répondre à cette question, nous avons utilisé des algorithmes de découverte de sous-groupes et mesuré la contribution de chaque individu à partir des motifs extraits.

A. La découverte de sous-groupes : un outil permettant de faire le lien entre l'olfaction et les émotions

La découverte de sous-groupes et plus généralement la découverte de modèles exceptionnels permet d'extraire des motifs qui regroupent des entités pour lesquelles le modèle induit sur les attributs de classe est significativement différent du modèle induit par l'ensemble des objets du jeu de données. Ces approches nous ont particulièrement intéressés car elles permettent d'extraire des modèles locaux et interprétables. L'intelligibilité des modèles était importante dans le cadre de ce travail de thèse pour pouvoir mieux comprendre la neurophysiologie de l'émotion. Par ailleurs, pour prendre en compte la variabilité individuelle, l'extraction de modèles à une échelle locale plutôt que sur l'ensemble des données nous semblait intéressante. Ainsi ces méthodes nous ont paru être de bonnes candidates pour analyser nos données neurophysiologiques.

Différentes approches existent dans ce domaine. Pour l'analyse de l'activité du système nerveux autonome (étude 1), nous nous sommes tournés vers une approche heuristique car les algorithmes exacts testés ne parvenaient pas à explorer les sept dimensions d'attributs physiologiques réels en un temps raisonnable. Pour cela nous avons utilisé la recherche arborescente de Monte-Carlo, permettant d'élaguer l'espace de recherche rapidement et de donner une réponse dans un temps imparti. Ce type de recherche a déjà été utilisé avec succès en olfaction pour faire le lien entre structures physicochimiques de la molécule odorante (ex. poids moléculaire, nombre d'atomes etc.) et qualités olfactives (ex. fruité, floral) [279–281]. Par rapport à ces analyses antérieures, un algorithme plus récent et performant a été sélectionné [294]. L'algorithme utilisé pour étudier l'activité du système nerveux central (étude 2 et 3) a été initialement développé dans le domaine de la géographie. Il a été utilisé afin de localiser et décrire les activités touristiques ou criminelles dans différents quartiers à partir de cartes géographiques [295]. Dans notre étude, nous avions pour objectif de localiser des activités spatiales à partir de données IRMf. Par ailleurs, la modélisation sous forme de graphe nous permettait d'inclure un lien entre les voxels adjacents et de créer une carte 3D de la région cérébrale étudiée. Cet algorithme nous a donc semblé en adéquation avec nos données et notre problématique. De plus, la découverte de sous-graphes exceptionnels avait précédemment été testée avec différentes modélisations pour localiser le traitement d'odeurs à partir de données IRMf [282]. La réflexion était similaire pour l'étude 3, dans laquelle la modélisation sous-forme de graphe a permis de relier les points temps-fréquences et de réaliser une recherche de motifs à partir de cartes EEG temps-fréquences.

En somme, en appliquant ces approches de découverte de sous-groupes à trois jeux de données, nous avons pu extraire :

- 1) Des règles décrivant les réponses émotionnelles physiologiques qui caractérisent des expériences subjectives particulières d'odeurs (études 1 et 4). Ces règles sont définies par des conjonctions de conditions sur des attributs physiologiques (température, fréquence cardiaque, fréquence respiratoire et conductance cutanée).
- 2) Des ensembles de voxels associés aux caractères plaisant ou déplaisant au sein de l'amygdale et du cortex piriforme, à partir de données IRMf (étude 2).
- 3) Des activités temps et fréquences caractéristiques de la valence hédonique et de l'éveil émotionnel ressentis lors de la perception d'une odeur, à partir de données EEG (étude 3).

Cette approche a mis en évidence que les odeurs agréables et désagréables agissent différemment sur l'activité du système nerveux périphérique et central. D'un point de vue de la modélisation descriptive, des résultats originaux ont été observés dans les trois analyses.

Nous avons montré la pertinence des résultats de deux manières, soit en réalisant une étude classique pour comparer, soit en comparant directement avec la littérature :

- 1) Par exemple, les résultats de l'étude 1 montrent, entre autres, une diminution de l'activité cardiaque pour les odeurs désagréables, une réponse électrodermale qui augmente avec l'intensité de l'odeur ainsi qu'une activité électrodermale forte pour les odeurs déplaisantes. Ces données soutiennent l'hypothèse d'une réponse somatique plus forte pour les stimuli olfactifs aversifs, un résultat déjà observé dans les études antérieures [195,249,251,296].
- 2) Les résultats de l'étude 2 révèlent une réponse aux odeurs désagréables latéralisée dans la partie antérieure gauche du cortex piriforme, zone dans laquelle la réponse hémodynamique était plus élevée pour les odeurs désagréables que pour les odeurs agréables [183,185]. Ce résultat est également confirmé par une analyse univariée menée en parallèle sur le même jeu de données.
- 3) Quant aux résultats de l'étude 3, une réaction cognitive tardive caractérise la perception d'odeurs plaisantes, et à l'inverse il existe une réaction très précoce pour les odeurs déplaisantes. De précédentes études utilisant des données d'électroencéphalographie [297], des activités olfactomotrices [298,299] et des temps de réaction [300,301] montrent également une réaction plus rapide pour les odeurs désagréables que pour les odeurs agréables.

Néanmoins nos analyses descriptives apportent une contribution complémentaire aux analyses utilisées jusqu'à présent et permettent, par ailleurs, de s'affranchir de certaines de leurs limites. Les outils couramment utilisés dans les analyses antérieures étudiant la neurophysiologie de l'émotion incluent entre autres les tests statistiques et des classifieurs. Les tests statistiques, utilisés depuis longtemps, permettent de rejeter ou d'accepter une hypothèse précédemment formulée en comparant des réponses affectives à travers des conditions (ex. agréable vs désagréable). Les résultats de ces tests décrivent souvent des orientations générales et impliquent le plus souvent une seule variable [251,302–305] (ex.

le rythme cardiaque diminue lorsque l'on perçoit une odeur agréable). Les classifieurs, de plus en plus utilisés, permettent de prédire la valence ou l'émotion ressentie à partir de données neurophysiologiques. Plusieurs études ont montré que la variabilité individuelle présente dans les données en sciences affectives limite les scores de prédiction. En effet, de meilleurs résultats sont obtenus lorsqu'un modèle est entraîné seulement sur le sujet testé plutôt que sur l'ensemble des sujets [256,306,307]. De plus, les modèles de deep learning généralement très performants nécessitent un jeu de données de grande taille afin d'atteindre de très bons niveaux d'apprentissage. Malheureusement, ces modèles sont encore peu adaptés aux expériences sur l'odorat pour au moins deux raisons : 1/ les expérimentations d'enregistrement du système nerveux en réponse à des odeurs en laboratoire sont longues et coûteuses ce qui restreint le volume de données collectées, 2/ en allongeant le temps d'expériences en olfaction, on s'expose au risque de fatigue et d'habituation olfactive.

Dans les analyses de fouille de données proposées, les hypothèses ne sont pas systématiquement formulées en amont de l'étude. Cependant, il est possible de formuler de nouvelles hypothèses jusqu'alors insoupçonnées à partir des résultats observés et les tester a posteriori à l'aide de méthodes statistiques ou prédictives. Les méthodes statistiques peuvent alors être utilisées pour confirmer ou rejeter ces hypothèses. Par ailleurs, une approche prédictive peut aussi permettre de tester les règles trouvées sur d'autres jeux de données indépendants et d'observer si les prédictions sont correctes ou non.

De plus, ces analyses descriptives permettent d'obtenir des résultats plus détaillés et quantifiés que des observations sous forme de tendances statistiques ou de taux de classification. C'est le cas, par exemple, pour les variations physiologiques qui sont souvent représentées sous forme de moyennes accompagnées d'écart-types [113,194,302,308–311]. Les motifs générés dans l'étude 1 ont permis de quantifier la réaction physiologique par des intervalles précis sur les paramètres dans leur ensemble et ce, pour un ressenti subjectif particulier (ex. plaisir, intensité).

Par ailleurs, les analyses descriptives peuvent également prendre en compte plusieurs dimensions neurophysiologiques conjointement. C'est le cas dans l'étude 1, où les paramètres psychophysologiques sont étudiés collectivement contrairement aux études antérieures qui traite bien souvent chaque paramètre séparément [194,302,309,311]. De même dans l'étude 3, la valence et l'éveil des odeurs ont été étudiés ensemble. De cette manière, les motifs trouvés pour le « plaisant » caractérisent à la fois les odeurs « plaisantes et calmantes » et les odeurs « plaisantes et éveillantes ».

Pour finir, de par leur caractère générique, les approches que nous avons développées peuvent prendre en compte des données hétérogènes. Ainsi, elles ont permis d'étudier des jeux de données avec des stimuli différents dans l'étude 2. Elles peuvent à la fois être utilisées sur des jeux de données de taille moyenne (centaines d'instances) comme ceux utilisés dans l'étude 2 et 3, mais également sur des données nombreuses (étude 1) avec des milliers voir des milliards d'instances.

B. La quantification de la participation des individus aux motifs : des mesures originales permettant de mieux comprendre la variabilité inter-individuelle

Comme nous l'avons évoqué tout au long de cette thèse, les analyses classiquement utilisées en neurophysiologie de l'olfaction ne prennent que très peu en compte la variabilité inter-individuelle. En effet, ces approches standards utilisent de nombreux types de tests statistiques (ex. Wilcoxon, Kruskal Wallis [302], ANOVA [308–311,194,312,313], MANOVA [113,314–317]) pour étudier la neurophysiologie de l'émotion et de l'olfaction. Les paramètres les plus couramment considérés sont des moyennes et des médianes de différents groupes, sans toujours s'intéresser à la variabilité inhérente aux différents jeux de données.

Les approches de science des données utilisées sont généralement des approches de classification qui tentent de chercher des modèles de discrimination entre différents états affectifs [252,253,256,306,307,318–329]. Trouver une règle universelle qui fonctionne sur tout le monde est intéressant mais ne rend cependant pas compte de la variabilité individuelle. D'ailleurs, des études de classification des émotions perçues à partir de l'activité du système nerveux périphérique [307] ou central [256,306,325], montrent une augmentation significative du score de prédiction lorsque le classifieur est appliqué par sujet plutôt que sur l'ensemble des sujets. Si la variabilité entrave la classification, c'est qu'elle a un impact significatif sur les données enregistrées et qu'il faut remettre l'individu au centre de l'analyse.

La prise en compte de la variabilité, omniprésente au niveau émotionnel et olfactif, représente une contribution importante dans nos analyses. Dans les trois études, nous avons pu quantifier l'implication de chaque individu dans les motifs générés en appliquant la mesure de qualité utilisée dans l'approche descriptive à l'échelle de l'individu ou d'un sous-groupe d'individus. Dans l'étude 1, nous nous sommes intéressés à la participation d'un individu : la mesure de qualité du motif était calculée uniquement sur les données de l'individu. Dans l'étude 2 et 3, la manière de calculer la participation des individus a évolué, et nous nous sommes intéressés à la contribution de l'individu dans le groupe en utilisant une approche provenant de la théorie des jeux appelée « Valeur de Shapley » [330]. Cette valeur était initialement utilisée dans le cadre de jeux coopératifs, lorsqu'une équipe collabore pour obtenir un gain. Cette métrique propose de répartir équitablement ce gain entre les joueurs en fonction de leur contribution. Ainsi nous avons trouvé pertinent de l'appliquer, de la même façon, sur les patterns afin de répartir équitablement le score de qualité du motif entre les individus en fonction de leur contribution. La quantification de la participation des individus, nous a permis de comparer la variabilité présente pour les odeurs agréables et désagréables dans les trois études. De cette façon, nous avons pu noter que selon la valence et l'éveil de l'émotion perçue, les réactions neurophysiologiques seront plus ou moins consensuelles entre individus (cf II.B). De plus, les valeurs de participation de populations différentes peuvent être comparées par des tests statistiques. Par ailleurs, dans l'étude 2, nous avons pu analyser si la participation aux motifs était dépendante de différents facteurs de variations comme l'âge (jeunes/âgés) et le sexe (hommes/femmes).

II. Représentation de la composante émotionnelle des odeurs au sein du système nerveux

Ainsi les approches de science des données développées permettent de faire le lien entre les émotions et l'olfaction, et sont complémentaires des analyses classiquement utilisées : elles permettent de générer de nouvelles hypothèses et d'étudier précisément la diversité inter-individuelle. Nous allons voir dans cette partie les principaux résultats neuroscientifiques qui en ressortent.

A. Émotions, odeurs et système nerveux

La représentation cérébrale du caractère hédonique des odeurs a été abordée à deux reprises dans cette thèse : dans l'étude 2, au niveau de régions olfactives primaires et dans l'étude 3, au niveau des électrodes frontales de l'EEG.

L'analyse de fouille de données dans l'étude 2 a révélé une activité asymétrique dans le cortex piriforme antérieur : les odeurs désagréables sont traitées préférentiellement à gauche alors que les odeurs agréables montrent des patterns d'activité dans le cortex piriforme antérieur droit. La latéralisation à gauche du traitement des odeurs désagréables au sein de cette structure a été confirmée par une analyse univariée réalisée sur ces mêmes données et par deux études antérieures. En effet, ces études décrivent une activité plus forte pour les odeurs désagréables que pour les odeurs agréables au niveau du cortex piriforme antérieur gauche [183,185]. Nos résultats montrent précisément (au voxel près) les zones cérébrales qui traitent le caractère agréable ou désagréable des odeurs au sein de cette structure. Concernant le cortex piriforme postérieur, les odeurs désagréables montrent également un pattern d'activité asymétrique gauche alors que des patterns bilatéraux apparaissent pour les odeurs plaisantes. Enfin, dans l'amygdale, des patterns d'activité bilatéraux sont observés pour les odeurs déplaisantes et plaisantes.

Dans l'étude 3, nous avons étudié les différences hémisphériques au niveau des électrodes frontales F3 et F4 de l'EEG. Cette étude a révélé une activité bilatérale symétrique lors de la perception de 3 catégories : odeurs éveillantes, odeurs calmantes et odeurs plaisantes. Pour la dimension d'éveil émotionnel (éveillant/calmant), l'analyse a révélé une réponse précoce et de relativement longue durée pour les odeurs éveillantes tandis que la perception d'odeurs calmantes est associée à une réaction tardive et de courte durée. Les odeurs éveillantes correspondent à des stimulations induisant une excitation émotionnelle forte signalant ainsi la présence d'un objet ou événement saillant dans l'environnement. De manière intéressante, une étude antérieure montre un temps de réaction plus rapide lors de la perception d'odeurs alimentaires vs. non alimentaires et une réaction plus rapide pour les odeurs désagréables vs. agréables [301]. Prises dans leur ensemble, ces études suggèrent que le système olfactif réagit plus rapidement aux stimuli écologiquement pertinents (ex. comestibles ou dangereux) afin de s'adapter au plus vite et décider d'un comportement d'approche ou d'évitement.

Pour les odeurs plaisantes, les patterns observés en F3 et F4 sont dans des intervalles de temps et de fréquences similaires et sont communs chez un très grand nombre d'individus. À l'inverse, pour la quatrième catégorie d'odeurs (c.-à-d odeurs déplaisantes), une activité asymétrique constituée par 4 patterns est observée : trois motifs sont présents en F4 (hémisphère droit) et un motif dans l'hémisphère gauche (F3). En F4, un pattern est précoce et associé à une activité alpha et deux patterns sont plus tardifs et associés à des fréquences plus hautes. Le pattern en F3 est caractérisé par une réponse tardive associée à des fréquences élevées (bêta et low-gamma). En termes temporels, les patterns tardifs en F3 et F4 n'apparaissent pas dans les mêmes fenêtres. On notera que la présence d'un motif à gauche lors de la perception d'odeurs déplaisantes n'est pas forcément en accord avec les études antérieures qui montrent une plus forte activité hémisphérique droite en réponse aux odeurs désagréables [306,325]. Cependant, le fait d'observer la majorité des motifs à droite au niveau du scalp dans notre analyse, et que ceux-ci regroupent au total la majorité des individus suggèrent néanmoins un traitement des odeurs désagréables plus accentué au sein de l'hémisphère droit, une asymétrie fonctionnelle connue pour être associée aux comportements de retrait [102,331].

La comparaison des études 2 et 3 révèle à première vue des contradictions. En effet, l'asymétrie observée lors de la perception d'odeurs déplaisantes n'implique pas le même hémisphère (cortex piriforme gauche pour l'étude 2 et électrode frontale droite pour l'étude 3). Cette comparaison inter-études reste néanmoins délicate dans la mesure où ce ne sont pas les mêmes méthodes qui ont été employées (IRMf dans l'étude 2 avec une très bonne résolution spatiale ; EEG dans l'étude 3 avec une résolution spatiale faible). Par ailleurs, les zones cérébrales explorées étaient également différentes : cortex piriforme pour l'étude 2, zone frontale superficielle pour l'étude 3. Ici, l'information la plus importante à retenir n'est pas forcément en lien avec l'hémisphère impliqué, mais serait plutôt qu'il semble exister une latéralisation à différents stades de traitement de l'information olfactive (plutôt à gauche pour le cortex piriforme, et peut être préférentiellement à droite pour des zones frontales). La vraie question ici est celle de savoir quelle est la fonction de ces asymétries ?

L'asymétrie fonctionnelle est un phénomène observé chez de nombreuses espèces animales [332]. Un exemple est la latéralisation manuelle observée chez les ours [333], chimpanzés [334,335] et humains [336] qui utilisent préférentiellement une main. Chez l'humain, cette latéralité manuelle se traduit par une latéralisation cérébrale contralatérale (du côté opposé à la main dominante).

Si la latéralisation fonctionnelle est aussi répandue, c'est certainement qu'elle a été acquise au fur et à mesure de l'évolution et qu'elle présente des avantages évolutifs. Deux avantages évolutifs de la latéralisation peuvent être considérés :

- 1) *La réalisation de tâches simultanées.* Par exemple, une latéralisation a été observée chez des individus parlant couramment l'anglais et le langage des signes. Les stimuli langagiers auditifs activaient l'hémisphère gauche, tandis que les stimuli langagiers visuels étaient traités dans l'hémisphère droit. De cette façon, les individus pouvaient suivre simultanément un discours parlé et signé [337].
- 2) *Une meilleure performance.* Par exemple, les chimpanzés sauvages ayant une préférence manuelle forte et donc une latéralisation forte, capturent plus de termites en une seule fois que les chimpanzés utilisant leurs deux mains [335]. De la même

façon, les seiches latéralisées sont plus rapides pour attaquer une proie que les seiches non latéralisées [338].

Ainsi, si nous faisons le lien avec la latéralisation en matière d'odeurs, celle-ci pourrait permettre de détecter par exemple des odeurs désagréables et réaliser dans le même temps d'autres tâches (fuite, approche, prise de conscience du stimulus, analyse de la situation etc.). Si l'on considère le deuxième avantage listé, on peut faire l'hypothèse que les odeurs désagréables sont latéralisées afin de déclencher une réaction plus performante ou plus rapide. Il est à noter que dans l'étude 3, nous observons justement une réaction plus rapide pour les odeurs désagréables que pour les odeurs agréables.

B. Variabilité interindividuelle, odeurs et système nerveux

Dans les trois études, pour tous les motifs générés, le niveau d'implication de chaque individu a été quantifié. Un niveau d'implication positif indiquait que l'individu participait au motif. À partir de ces informations, nous avons pu comparer les participations des motifs associés à un même état affectif (ex. plaisant) afin de caractériser la distribution d'individus dans ces différents motifs. Il s'agissait aussi de comprendre si les différents motifs observés étaient caractéristiques pour des personnes différentes et donc reflétaient une grande diversité.

De cette manière, les études 1 et 2 ont révélé une plus grande variabilité interindividuelle dans les réponses aux odeurs agréables et neutres par rapport aux odeurs désagréables. Ce résultat est en accord avec une étude utilisant une approche machine learning qui prédit correctement les réponses neurophysiologiques des odeurs déplaisantes, mais n'arrive pas à différencier les odeurs neutres et plaisantes [252]. Une réponse neurophysiologique quasiment universelle et stable (faiblement variable) pour les odeurs désagréables pourrait s'expliquer par un besoin commun de détecter les stimuli potentiellement dangereux et d'adopter une réaction adaptée [339].

À contrario, l'étude 3 semble montrer une réaction cérébrale plus homogène pour les odeurs agréables que désagréables. En effet, la perception d'odeurs plaisantes est caractérisée par une réponse bilatérale tardive dans les basses fréquences. Pour les odeurs déplaisantes, nous avons observé 4 types de réponses non exclusives. La réponse la plus consensuelle étant la réaction immédiate, on peut faire l'hypothèse qu'elle présente un avantage sélectif plus intéressant que les patterns apparaissant plus tardivement.

L'hypothèse précédemment formulée dans les études 1 et 2, postulant, que les odeurs désagréables créent une réaction plus universelle que les odeurs agréables, n'est donc pas vérifiée dans cette troisième étude. Pour expliquer cette contradiction apparente, nous pouvons émettre différentes hypothèses. La première hypothèse concerne la technique de fouille de données utilisée : lorsque de nombreux patterns sont trouvés, il y a statistiquement moins de chance que ce soient les mêmes individus qui participent aux différents motifs. Ainsi, plus il y a de motifs, plus il y a de diversité. Cette première hypothèse est contredite par le fait que nous observons autant de motifs pour les odeurs désagréables que pour les

odeurs éveillantes, alors que pour ces dernières, nous observons beaucoup d'individus communs dans les patterns observés.

La seconde hypothèse est qu'il y aurait une réponse neurophysiologique commune pour les odeurs désagréables mais qu'elle se passe à un instant différent selon les individus. Les résultats des études 1 et 2 rendent compte d'une représentation multidimensionnelle et/ou spatiale du traitement des odeurs. De manière complémentaire, les résultats de l'étude 3, apportent un éclairage sur le traitement temporel de ces mêmes stimulations olfactives. Par ailleurs, dans l'étude 1, on peut tout de même remarquer que les deux attributs temporels (latence et rise time) sont caractérisés sur des plages temporelles plus longues pour les odeurs déplaisantes (2,87 sec pour la latence et 6,81 sec pour le rise time) que les odeurs plaisantes (0,49 sec pour la latence et 2,61 sec pour le rise time). Ainsi, une diversité interindividuelle est peut-être présente dans ces longs laps de temps et il serait intéressant de l'étudier.

Ainsi au niveau temporel, la perception d'odeurs agréables est associée à une réaction cognitive tardive. Cependant, ces odeurs n'activent pas les mêmes régions du cortex piriforme et de l'amygdale et ne créent pas les mêmes réactions physiologiques chez tout le monde. Concernant les odeurs désagréables, peu de variabilité est constatée au niveau de la conductance cutanée et des régions cérébrales (amygdale et piriforme) mais ces activations similaires ne se passent pas forcément au même moment chez tous les individus.

Suite à ces résultats, on peut s'interroger sur le rôle de la variabilité interindividuelle en matière de perception hédonique des odeurs. La variabilité olfactive est le résultat d'une réponse adaptative. Elle se présente de façon intra et inter-individuelle et exprime l'exposition aux odorants à laquelle chacun fait face mais peut aussi être liée à des éléments d'ordre génétique. Ainsi nous avons acquis une aversion ou un attrait pour les odeurs en fonction de leurs pertinences dans notre environnement. Le système olfactif s'adapte à l'échelle d'une population mais il s'adapte également à l'échelle d'un individu. En effet, durant la vie, une variabilité intra-individuelle est présente en fonction des apprentissages et des expériences vécues afin d'éviter de répéter une expérience désagréable et afin de retrouver une expérience agréable. À l'échelle d'une année, la variabilité olfactive permet, par exemple, à une femme enceinte de protéger son fœtus à l'aide d'un odorat exacerbé durant le premier trimestre de la grossesse [215,216]. De même, à l'échelle d'un mois, la variabilité permet à une femme d'avoir un odorat plus développé lorsqu'elle est en phase fertile [213,214]. À l'échelle d'une journée, la variabilité intra-individuelle permet de réguler notre appétit [244].

Ainsi, si nous trouvons des caractéristiques universelles dans le caractère plaisant ou déplaisant d'une odeur, c'est probablement que celles-ci ont un avantage évolutif. Le caractère déplaisant a un intérêt évolutif fort puisqu'il permet d'éviter des produits toxiques ou des aliments avariés. On peut supposer qu'il y a moins de variabilité au niveau physiologique et cérébral pour les odeurs désagréables car, il faut s'apprêter à avoir une réaction de fuite vis-à-vis de la source. L'étude 3 montre une diversité dans le traitement temporel de ces informations olfactives désagréables. Il est alors possible de penser que même si le réseau (spatial) olfactif est commun aux individus, son utilisation dans le temps est variable. La question de savoir pourquoi cette diversité temporelle existe reste encore à explorer. Ici, plusieurs facteurs de variations seront à considérer incluant par exemple l'échantillonnage des odeurs au niveau de la cavité nasale (activité olfactomotrice) ou encore

des différences physiologiques et/ou démographiques entre individus (état hormonal, sexe, âge, etc). À l'inverse, pour les odeurs agréables, on peut faire l'hypothèse du recrutement de différents réseaux cérébraux entre les individus (expliquant ainsi les données des études 1 et 2). L'implication de ces réseaux pourrait néanmoins impliquer un mécanisme temporel commun, ce qui expliquerait la forte concordance entre individus des patterns EEG générés dans l'étude 3.

III. Limites

Les études présentées tout au long de ce manuscrit apportent de nouvelles contributions et sont prometteuses. Cependant, les méthodologies développées peuvent encore être améliorées avec des données expérimentales plus appropriées. Ces limites expérimentales et méthodologiques sont présentées ci-après avec des pistes d'amélioration possibles.

A. Limites expérimentales

1. Nombre d'individus et d'odeurs

Les données des trois études sont caractérisées par un grand volume de données. En effet, celles-ci comportent de nombreuses dimensions caractérisant les enregistrements du système nerveux périphérique et central. Cependant, les trois jeux de données contenaient, dans l'ensemble, peu d'individus et de stimuli (odorants) ce qui a pu biaiser les résultats présentés et limiter par conséquent l'exploration de la variabilité individuelle en matière d'odeurs. Dans notre première étude, 109 odorants et uniquement 22 individus ont été étudiés. Dans la seconde, il y avait 42 individus et simplement 12 odorants. Enfin, dans la troisième et dernière étude, seulement 6 odorants et 24 participants ont été intégrés au protocole.

Faire des études avec un grand nombre de participants est très coûteux (en temps et financièrement). Par ailleurs, l'utilisation de nombreux odorants peut rallonger le temps d'expérimentation. En réalisant des sessions expérimentales longues, on s'expose au risque de fatigue de la part du participant et à des effets d'habituation aux odorants [340]. Parmi les solutions envisagées pour parer à ce problème, on peut citer :

- 1) *Réaliser une expérience avec beaucoup d'individus qui sentent chacun des odeurs différentes.* Le nombre d'individus étant élevé, nous pourrions étudier différents facteurs de variabilité. Le nombre d'odeurs présentées au sujet ne serait pas trop élevé afin de réaliser une expérimentation qui soit faisable en une séance. Chaque individu sentirait un ensemble spécifique d'odeurs. Vu que nous intéressons ici à la perception émotionnelle d'un stimulus, nous n'avons pas forcément besoin que celui-ci soit le même pour tout le monde. Plus le nombre d'odeurs utilisées est important, plus nous sommes sûrs de capturer un phénomène lié à la perception émotionnelle et non pas à d'autres propriétés des odorants comme la chimie ou la perception trigéminal. Si par exemple, dans le jeu de données, seules les odeurs soufrées ont

été jugées désagréables, nous ne pouvons pas savoir si les motifs découverts pour la perception des odeurs désagréable correspondent au « désagréable » ou au « soufre ». Par conséquent, il est essentiel d'étudier une grande diversité d'odeurs.

- 2) *Faire une méta-analyse à partir de données venant d'expériences similaires.* De la même manière que dans l'étude 2 où l'on a combiné deux expérimentations similaires avec des odeurs différentes, nous pourrions envisager de combiner des jeux de données provenant d'expérimentations différentes. Ainsi nous pourrions, étudier un grand nombre d'individus et d'odeurs et exploiter de nouveaux facteurs de variation. Par exemple, si une expérience similaire se déroule dans le monde entier, nous pourrions étudier des facteurs environnementaux et culturels.

2. *Facteurs de variabilité*

Dans les différentes études présentées, nous avons étudié les facteurs de variabilité à notre disposition comme l'âge et le sexe. Aucun motif significatif n'est ressorti de ces analyses. Ce ne sont apparemment pas des conditions suffisantes et il faudrait pouvoir également prendre en compte les multiples autres facteurs, présentés dans la partie II.D de l'introduction, comme la phase hormonale, l'appétit, la génétique, l'apprentissage etc. Les méthodes que nous avons développées n'ont pas été appliquées à des données comprenant ces informations, mais permettent d'étudier de la même façon d'autres facteurs de variation. Il conviendra néanmoins d'avoir un échantillon représentatif pour chaque catégorie d'intérêt et donc un assez grand nombre d'individus.

3. *Données utilisées*

En continuité avec les limites précédentes, il serait pertinent d'acquérir des données supplémentaires concernant les participants. Par exemple, la mise en place de test de profilage génétique [341] pour connaître les récepteurs olfactifs fonctionnels présents chez l'individu pourraient permettre de comprendre si certaines dimensions de la perception olfactive peuvent être mises en lien avec des facteurs génétiques.

De plus, il peut être également intéressant d'enregistrer d'autres paramètres physiologiques encore plus discriminants dans la classification des réponses affectives comme la variabilité du rythme cardiaque qui peut être mesurée à l'aide de capteurs ECG. Par ailleurs, les trois expériences analysées étaient complètement indépendantes : des odeurs communes ont été utilisées mais les participants des expériences étaient différents entre chaque étude. Il serait intéressant de pouvoir créer des expériences qui combinent des enregistrements différents du système nerveux afin de pouvoir faire le lien entre les différents types de données (physiologiques, IRMf et EEG). Cela peut permettre de comprendre si les individus qui ont une réaction émotionnelle similaire au niveau physiologique ont également une réaction similaire au niveau cérébral, par exemple. Cette combinaison peut se faire de différentes manières :

- 1) Soit l'expérience est découpée en trois temps et les mêmes individus suivent les trois enregistrements séparément mais dans des conditions identiques.

- 2) Soit les trois enregistrements sont effectués en simultanée grâce à des technologies d'EEG [342,343] et d'enregistrements physiologiques [344,345] adaptés à l'IRMf. Cette expérience serait lourde et coûteuse à mettre en place mais cela donnerait accès à des informations multiples pour une même perception. Ces informations différentes se complèteraient et la relation entre les différentes dimensions pourrait être étudiée.

B. Limites méthodologiques

Nous pouvons également remarquer que les approches élaborées apportent une contribution intéressante mais présentent encore quelques limites. Dans les approches descriptives, il est souvent difficile de trouver le bon niveau de description pour capturer un phénomène. En effet, un phénomène peut être décrit de multiples façons, il faut ainsi trouver une description avec un niveau de détails suffisant et qui reste compréhensible et interprétable. Ainsi dans l'étude psychophysique, par exemple, seules les règles à une dimension ont été commentées, les autres étaient beaucoup plus précises mais beaucoup plus difficilement interprétables par les experts du domaine. Ainsi, il est certainement possible de trouver des langages de description plus adaptés que ceux que nous avons choisis : une nouvelle façon de décrire les données sans tomber dans la sur-description ou dans la description de motifs inintéressants car triviaux.

Par ailleurs, les facteurs de variation pris en compte ont été étudiés indépendamment les uns des autres. Il peut être pertinent d'étudier les combinaisons de facteurs. Ainsi dans l'étude 2, même si rien de significatif n'est trouvé pour les hommes, les femmes, les jeunes et les personnes âgées, il pourrait y avoir des motifs concernant, par exemple, seulement les femmes âgées. L'exploitation des participations individuelles peut être étudiée différemment afin de rechercher une description précise des participants d'un pattern en fonction de tous les facteurs à notre disposition.

IV. Perspectives

Dans cette partie nous aborderons les perspectives possibles de ce travail. Des applications potentielles des approches et des résultats générés seront proposées. Ici, les méthodologies proposées pourraient être envisagées dans un cadre plus large que la seule compréhension du lien entre l'olfaction et les émotions. Ensuite, nous décrirons des impacts possibles de la compréhension de la neurophysiologie des émotions et de l'olfaction sur la société.

A. Approches développées prometteuses et polyvalentes

Les approches développées dans les différentes études ne sont pas exclusives au jeu de données utilisé. Elles peuvent tout à fait être utilisées dans d'autres contextes neuroscientifiques que l'étude de l'olfaction, pour caractériser d'autres attributs cibles que le caractère hédonique ou l'intensité et pour tester d'autres facteurs de variation tels que

l'âge et le sexe. Par ailleurs, le tutoriel présenté dans l'étude 5 est appliqué à un tableau de données de conductance cutanée, mais il pourrait très bien être utilisé sur d'autres données numériques comme d'autres paramètres physiologiques ou des mesures comportementales et d'expressions faciales par exemple.

Par ailleurs, ces approches peuvent prendre en compte un gros volume de données, ce qui peut s'avérer être particulièrement pertinent dans le cas d'analyse de questionnaires ou d'expériences écologiques réalisés par de nombreux participants à leur domicile (voir étude 5).

De plus, ces méthodes ne nécessitent pas que les individus étudiés aient tous perçu les mêmes stimuli. Elles peuvent ainsi être utilisées sur plusieurs jeux de données en même temps pour effectuer une méta-analyse.

B. La compréhension du lien odeurs et émotions : créatrice de nouvelles opportunités

Améliorer notre compréhension de la neurophysiologie de l'émotion dans la perception olfactive pourrait avoir des conséquences bénéfiques pour la société.

Mieux comprendre le fonctionnement de l'odorat et des émotions est nécessaire en santé pour pouvoir appréhender au mieux les dysfonctionnements par la suite. Cette méthodologie pourrait également être utilisée pour étudier certains déficits sensoriels tels que des hyposmies post-infectieuses ou des maladies neurodégénératives comme la maladie d'Alzheimer. Des différences de participation entre sujets sains et sujets malades pourraient émerger mais également des différences entre différents groupes d'individus au sein d'une même maladie. Ainsi ces études pourraient aider à la recherche d'un traitement global mais également à un traitement spécifique par sous-population, qui serait par conséquent plus adapté et plus efficace.

D'autres domaines d'applications futures peuvent également être envisagés, notamment dans les secteurs de l'alimentaire et dans l'industrie du parfum toujours à la recherche de nouvelles saveurs et de fragrances plaisantes. L'industrie du jeu et de l'art pourrait également être intéressée par les effets émotionnels des odeurs pour augmenter l'immersion dans une œuvre [346]. En effet, on peut facilement imaginer utiliser la diffusion d'odeurs dans un cadre artistique afin que l'œuvre suscite plus fortement une émotion : un tableau zen associé à une odeur apaisante ou encore un film comique diffusant une odeur joviale. En effet, on cherche de plus en plus à donner une dimension réaliste aux expériences cinématographiques [347] en utilisant des images 4k ou 3D, des sièges dynamiques, du son multicanal (utilisation de différentes sources sonores afin de créer un champ sonore de 360°) ; l'ajout d'odeurs suscitant des émotions sera peut-être courant dans les années à venir.

Le domaine du bien-être peut également se sentir concerné où l'on cherche à apaiser une douleur par l'inhalation d'odeurs [348]. Des objets innovants pourraient être développés dans ce sens. On pourrait, par exemple, concevoir un collier libérant des odeurs qui modulent les émotions négatives d'une personne (détectées par une montre connectée).

V. Le mot de la fin

Dans l'ensemble, les résultats de cette thèse montrent une utilisation pertinente des approches descriptives provenant de la science des données dans la compréhension de la neurophysiologie de l'émotion et de l'odorat et de sa grande variabilité inter-individuelle. En effet, nous avons montré que les résultats étaient d'une part en adéquation avec la littérature et, d'autre part, novateurs car ils apportent des éléments nouveaux par rapport aux études réalisées précédemment. Les approches descriptives ont ainsi un grand potentiel dans les recherches futures sur les émotions, l'olfaction et les neurosciences en général.

RÉFÉRENCES

- [1] Cairns D. *A Cultural History of the Emotions in Antiquity*. Bloomsbury Publishing; 2020.
- [2] Konstan D. Emotions and Morality: The View from Classical Antiquity. *Topoi*. 2015;34:401–407.
- [3] Renaut O. *Platon la Médiation des Émotions: L'Éducation du Thymos Dans les Dialogues*. Vrin; 2014.
- [4] Frère J. *Philosophie des émotions: les sages nous aident à en faire bon usage*. Editions Eyrolles; 2009.
- [5] Irons D. Descartes and Modern Theories of Emotion. *Philos Rev*. 1895;4:291–302.
- [6] Damasio A. Descartes' error: Emotion, Rationality and the Human Brain. N Y Putnam. 1994;352.
- [7] Darwin C, Prodger P. *The Expression of the Emotions in Man and Animals*. Oxford University Press; 1998.
- [8] Greenberg LS. *Emotion in Psychotherapy*. Guilford Press; 1990.
- [9] Lewis H. *Freud and Modern Psychology: Volume 1: The Emotional Basis of Mental Illness*. Springer Science & Business Media; 2013.
- [10] James W. Discussion: The physical basis of emotion. *Psychol Rev*. 1894;1:516–529.
- [11] James W. *The emotions*. Baltimore, MD, US: Williams & Wilkins Co; 1922. p. 135.
- [12] Cannon WB. Again the James-Lange and the thalamic theories of emotion. *Psychol Rev*. 1931;38:281–295.
- [13] Luminet O. *Psychologie des émotions: Confrontation et évitement*. De Boeck Supérieur; 2008.
- [14] *Dictionnaire*. LeRobert; 2022.
- [15] Buss AH, Plomin R. *Temperament (PLE: Emotion): Early Developing Personality Traits*. Psychology Press; 2014.
- [16] Watson D. *Mood and Temperament*. Guilford Press; 2000.
- [17] Malik F, Marwaha R. *Developmental Stages of Social Emotional Development In Children*. StatPearls StatPearls Publ Treasure Isl FL. 2022;
- [18] Scherer KR, Ekman P. *Approaches To Emotion*. Psychology Press; 1984.
- [19] Dalgleish T, Power M. *Handbook of Cognition and Emotion*. John Wiley & Sons; 2000.
- [20] Campbell N, Reece J, Cain M, et al. *Biologie*. 9th ed. Pearson; 2012.
- [21] Lange CG. The emotions. In: Haupt IA, editor. *Emot Vol 1*. Baltimore, MD, US: Williams & Wilkins Co; 1922. p. 33–90.
- [22] Bard P. On emotional expression after decortication with some remarks on certain theoretical views: Part I. *Psychol Rev*. 1934;41:309–329.
- [23] Schachter S, Singer J. Cognitive, social, and physiological determinants of emotional state. *Psychol Rev*. 1962;69:379–399.
- [24] Schachter S. The Interaction of Cognitive and Physiological Determinants of Emotional State. Much of the research described in this paper was supported by Grant MH 05203 from the National Institute of Mental Health, United States Public Health Service, and by Grant G 23758 from the National Science Foundation. In: Berkowitz L, editor. *Adv Exp Soc Psychol* [Internet]. Academic Press; 1964 [cited 2022 Oct 26]. p. 49–80. Available from: <https://www.sciencedirect.com/science/article/pii/S0065260108600489>.
- [25] Lazarus RS, Smith CA. Knowledge and Appraisal in the Cognition—Emotion Relationship. *Cogn Emot*. 1988;2:281–300.
- [26] Sander D, Grandjean D, Scherer KR. A systems approach to appraisal mechanisms in emotion. *Neural Netw*. 2005;18:317–352.

- [27] Siemer M, Mauss I, Gross JJ. Same situation--Different emotions: How appraisals shape our emotions. *Emotion*. 2007;7:592–600.
- [28] Christophe V. *Les Émotions: Tour d’horizon des principales théories*. Presses Univ. Septentrion; 1998.
- [29] Coppin G, Sander D. Théories et concepts contemporains en psychologie de l’émotion [Internet]. *Systèmes Interact. Émotionnelle*. Hermès Science publications-Lavoisier; 2010 [cited 2022 Aug 29]. p. 25. Available from: <https://archive-ouverte.unige.ch/unige:34368>.
- [30] Matsumoto D, Hwang HSC. Culture, Emotion, and Expression. *Cross-Cult Psychol* [Internet]. John Wiley & Sons, Ltd; 2019 [cited 2022 Oct 26]. p. 501–515. Available from: <https://onlinelibrary.wiley.com/doi/abs/10.1002/9781119519348.ch24>.
- [31] Matsumoto D, Ekman P. American-Japanese cultural differences in intensity ratings of facial expressions of emotion. *Motiv Emot*. 1989;13:143–157.
- [32] Plutchik R. Chapter 1 - A GENERAL PSYCHOEVOLUTIONARY THEORY OF EMOTION. In: Plutchik R, Kellerman H, editors. *Theor Emot* [Internet]. Academic Press; 1980 [cited 2022 Oct 26]. p. 3–33. Available from: <https://www.sciencedirect.com/science/article/pii/B9780125587013500077>.
- [33] Barrett LF, Russell JA. The Structure of Current Affect: Controversies and Emerging Consensus. *Curr Dir Psychol Sci*. 1999;8:10–14.
- [34] Frijda NH. *The Emotions*. Cambridge University Press; 1986.
- [35] Scherer KR. Emotion as a multicomponent process: A model and some cross-cultural data. *Rev Personal Soc Psychol*. 1984;5:37–63.
- [36] Izard CE, Kagan J, Zajonc RB. *Emotions, Cognition, and Behavior*. CUP Archive; 1984.
- [37] Ahmed W, van der Werf G, Minnaert A, et al. Students’ daily emotions in the classroom: Intra-individual variability and appraisal correlates. *Br J Educ Psychol*. 2010;80:583–597.
- [38] Scherer KR, Schorr A, Johnstone T. *Appraisal Processes in Emotion: Theory, Methods, Research*. Oxford University Press; 2001.
- [39] Frijda NH, Kuipers P, ter Schure E. Relations among emotion, appraisal, and emotional action readiness. *J Pers Soc Psychol*. 1989;57:212–228.
- [40] Richelle M, Requin J, Robert M, editors. *Traité de psychologie expérimentale*. Paris, France: Presses universitaires de France; 1994.
- [41] Bänziger T, Grandjean D, Bernard PJ, et al. Prosodie de l’émotion : étude de l’encodage et du décodage. :27.
- [42] Bergmann G. Stuttering as a Prosodie Disturbance: A Link between Speech Execution and Emotional Processes. In: Peters HFM, Hulstijn W, editors. *Speech Mot Dyn Stuttering* [Internet]. Vienna: Springer; 1987 [cited 2022 Aug 31]. p. 393–407. Available from: https://doi.org/10.1007/978-3-7091-6969-8_31.
- [43] Jouanne C. L’alexithymie : entre déficit émotionnel et processus adaptatif. *Psychotropes*. 2006;12:193–209.
- [44] Doba K, Nandrino J-L. Dérèglements émotionnels dans l’anorexie mentale : perspective interactionniste. *Bull Psychol*. 2008;Numéro 494:145–157.
- [45] Purves D, editor. *Neuroscience*. 3rd ed. Sunderland, Mass: Sinauer Associates, Publishers; 2004.
- [46] Pasquier A. Les troubles de l’émotion. *Santé Ment*. 2013;32.
- [47] Auffret M. “Stranger than fiction”: the nineteenth-century media coverage of the case of Phineas Gage (P10-7.003). *Neurology* [Internet]. 2022 [cited 2022 Oct 27];98. Available from: https://n.neurology.org/content/98/18_Supplement/1560.
- [48] O’Driscoll K, Leach JP. “No longer Gage”: an iron bar through the head. *Early*

observations of personality change after injury to the prefrontal cortex. *BMJ*. 1998;317:1673–1674.

[49] Damasio H, Grabowski T, Frank R, et al. The Return of Phineas Gage: Clues About the Brain from the Skull of a Famous Patient. *Science*. 1994;264:1102–1105.

[50] Cato MA, Delis DC, Abildskov TJ, et al. Assessing the elusive cognitive deficits associated with ventromedial prefrontal damage: A case of a modern-day Phineas Gage. *J Int Neuropsychol Soc*. 2004;10:453–465.

[51] Dimitrov M, Phipps M, Zahn TP, et al. A thoroughly modern gage. *Neurocase*. 1999;5:345–354.

[52] Goel V, Grafman J, Tajik J, et al. A study of the performance of patients with frontal lobe lesions in a financial planning task. *Brain*. 1997;120:1805–1822.

[53] Hornak J, Rolls ET, Wade D. Face and voice expression identification in patients with emotional and behavioural changes following ventral frontal lobe damage. *Neuropsychologia*. 1996;34:247–261.

[54] Chow TW. Personality in frontal lobe disorders. *Curr Psychiatry Rep*. 2000;2:446–451.

[55] Keltner D, Ekman P, Kazdin A. *Encyclopedia of Psychology*. Oxford University Press; 2000.

[56] Plutchik R. What is an Emotion? *J Psychol*. 1965;61:295–303.

[57] Durbano F. *Anxiety: An Adaptive Emotion. New Insights Anxiety Disord. BoD – Books on Demand*; 2013.

[58] Avier G. L'émotion. Folle du logis ou ancre de carrière, l'émotion vectrice de sens chez le manager public dans les organisations publiques territoriales. *Rev Int Psychosociologie Gest Comport Organ*. 2013;XIX:43–71.

[59] Lopes PN, Salovey P, Côté S, et al. Emotion Regulation Abilities and the Quality of Social Interaction. Petty RE, editor. *Emotion*. 2005;5:113–118.

[60] Colletta J-M. *Les émotions: cognition, langage et développement*. Editions Mardaga; 2003.

[61] Nakahashi W, Ohtsuki H. When is emotional contagion adaptive? *J Theor Biol*. 2015;380:480–488.

[62] Tyng CM, Amin HU, Saad MNM, et al. The Influences of Emotion on Learning and Memory. *Front Psychol* [Internet]. 2017 [cited 2022 Oct 27];8. Available from: <https://www.frontiersin.org/articles/10.3389/fpsyg.2017.01454>.

[63] Meneses A, Liy-Salmeron G. Serotonin and emotion, learning and memory. *Rev Neurosci*. 2012;23:543–553.

[64] Martinot J-L, Zilbovicius M, Paillère M-L, et al. Développement cérébral en psychiatrie de l'enfant et de l'adolescent. Approche par l'imagerie. *L'Évolution Psychiatr*. 2002;67:122–154.

[65] Campos JJ. A new perspective on emotions. *Child Abuse Negl*. 1984;8:147–156.

[66] Rochat P. Comment la honte vient aux enfants. *Communications*. 2013;93:69–84.

[67] Wierzbicka A. *Emotion, language, and cultural scripts. Emot Cult Empir Stud Mutual Infl.* Washington, DC, US: American Psychological Association; 1994. p. 133–196.

[68] Sauter DA, LeGuen O, Haun D. Categorical perception of emotional facial expressions does not require lexical categories. - *PsycNET. Emotion*. 2011;11:1479.

[69] Breugelmans SM, Poortinga YH. Emotion without a word: Shame and guilt among Rarámuri Indians and rural Javanese. *J Pers Soc Psychol*. 2006;91:1111–1122.

[70] Lomas T. *Translating Happiness: A Cross-Cultural Lexicon of Well-Being*. MIT Press; 2019.

[71] Lomas T. *The Happiness Dictionary: Words from Around the World to Help Us Lead a Richer Life*. Hachette UK; 2018.

- [72] Russell JA. Culture and the categorization of emotions. *Psychol Bull.* 1991;110:426–450.
- [73] Ekman P, Levenson RW, Friesen WV. Autonomic Nervous System Activity Distinguishes Among Emotions. *Science.* 1983;221:1208–1210.
- [74] Kreibig SD. Autonomic nervous system activity in emotion: A review. *Biol Psychol.* 2010;84:394–421.
- [75] Kohler CG, Turner T, Stolar NM, et al. Differences in facial expressions of four universal emotions. *Psychiatry Res.* 2004;128:235–244.
- [76] Ekman P. Universals and cultural differences in facial expressions of emotion. *Nebr Symp Motiv.* 1971;19:207–283.
- [77] Izard CE. Innate and universal facial expressions: Evidence from developmental and cross-cultural research. *Psychol Bull.* 1994;115:288–299.
- [78] Ekman P. Facial Expression. *Handb Cogn Emot.* John Wiley & Sons; 2000.
- [79] Parr LA, Waller BM. Understanding chimpanzee facial expression: insights into the evolution of communication. *Soc Cogn Affect Neurosci.* 2006;1:221–228.
- [80] Boissy A, Arnould C, Chaillou E, et al. Emotions and cognition: a new approach to animal welfare. *Anim Welf.* 2007;16:37–43.
- [81] Lezama-García K, Orihuela A, Olmos-Hernández A, et al. Facial expressions and emotions in domestic animals. *CAB Rev.* 2019;14:1–12.
- [82] Briefer EF. Vocal expression of emotions in mammals: mechanisms of production and evidence. *J Zool.* 2012;288:1–20.
- [83] Paul ES, Harding EJ, Mendl M. Measuring emotional processes in animals: the utility of a cognitive approach. *Neurosci Biobehav Rev.* 2005;29:469–491.
- [84] Dantzer R. Can farm animal welfare be understood without taking into account the issues of emotion and cognition? *J Anim Sci.* 2002;80:E1–E9.
- [85] Désiré L, Boissy A, Veissier I, et al. A cognitive approach to emotions: are the responses to suddenness and novelty different? [Internet]. 2002 [cited 2022 Oct 30]. Available from: <https://hal.archives-ouvertes.fr/hal-02005079>.
- [86] de Waal FBM. What is an animal emotion? *Ann N Y Acad Sci.* 2011;1224:191–206.
- [87] Smith AV, Proops L, Grounds K, et al. Functionally relevant responses to human facial expressions of emotion in the domestic horse (*Equus caballus*). *Biol Lett.* 2016;12:20150907.
- [88] Semin GR, Scandurra A, Baragli P, et al. Inter- and Intra-Species Communication of Emotion: Chemosignals as the Neglected Medium. *Animals.* 2019;9:887.
- [89] D’Aniello B, Semin GR, Alterisio A, et al. Interspecies transmission of emotional information via chemosignals: from humans to dogs (*Canis lupus familiaris*). *Anim Cogn.* 2018;21:67–78.
- [90] PAPEZ JW. A PROPOSED MECHANISM OF EMOTION. *Arch Neurol Psychiatry.* 1937;38:725–743.
- [91] MacLEAN PD. THE LIMBIC SYSTEM (“VISCERAL BRAIN”) AND EMOTIONAL BEHAVIOR. *AMA Arch Neurol Psychiatry.* 1955;73:130–134.
- [92] LeDoux JE. Emotion and the amygdala. *Amygdala Neurobiol Asp Emot Mem Ment Dysfunct.* New York, NY, US: Wiley-Liss; 1992. p. 339–351.
- [93] Hariri AR, Mattay VS, Tessitore A, et al. Neocortical modulation of the amygdala response to fearful stimuli. *Biol Psychiatry.* 2003;53:494–501.
- [94] Davidson RJ. Hemispheric asymmetry and emotion. *Approaches to emotion. Approaches Emot.* Psychology Press; 1984.
- [95] Silberman EK, Weingartner H. Hemispheric lateralization of functions related to emotion. *Brain Cogn.* 1986;5:322–353.
- [96] Black FW. Unilateral Brain Lesions and MMPI Performance: A Preliminary Study.

Percept Mot Skills. 1975;40:87–93.

[97] Gasparini WG, Satz P, Heilman K, et al. Hemispheric asymmetries of affective processing as determined by the Minnesota Multiphasic Personality Inventory. *J Neurol Neurosurg Psychiatry*. 1978;41:470–473.

[98] Goldstein K. *The organism: A holistic approach to biology derived from pathological data in man*. Brooklyn, NY, US: Zone Books; 1995. p. 422.

[99] Gainotti G. Emotional Behavior and Hemispheric Side of the Lesion. *Cortex*. 1972;8:41–55.

[100] Robinson RG, Price TR. Post-stroke depressive disorders: a follow-up study of 103 patients. *Stroke*. 1982;13:635–641.

[101] Sackeim HA, Greenberg MS, Weiman AL, et al. Hemispheric Asymmetry in the Expression of Positive and Negative Emotions: Neurologic Evidence. *Arch Neurol*. 1982;39:210–218.

[102] Harmon-Jones E, Gable PA. On the role of asymmetric frontal cortical activity in approach and withdrawal motivation: An updated review of the evidence. *Psychophysiology*. 2018;55:e12879.

[103] Ross ED, Mesulam M-M. Dominant Language Functions of the Right Hemisphere?: Prosody and Emotional Gesturing. *Arch Neurol*. 1979;36:144–148.

[104] Campbell R. The Lateralisation of Emotion: A Critical Review. *Int J Psychol*. 1982;17:211–229.

[105] Mandal MK, Ambady N. Laterality of facial expressions of emotion: Universal and culture-specific influences. *Behav Neurol*. 2004;15:23–34.

[106] Korb S, Delplanque S. Chapitre 8. Électromyographie (EMG). *Electrophysiol Cogn* [Internet]. Paris: Dunod; 2013 [cited 2022 Aug 30]. p. 179–194. Available from: <https://www.cairn.info/electrophysiologie-de-la-cognition--9782100593064-p-179.htm>.

[107] Bettadapura V. Face Expression Recognition and Analysis: The State of the Art [Internet]. arXiv; 2012 [cited 2022 Aug 30]. Available from: <http://arxiv.org/abs/1203.6722>.

[108] Bänziger T, Grandjean D, Scherer KR. Emotion recognition from expressions in face, voice, and body: The Multimodal Emotion Recognition Test (MERT). *Emotion*. 2009;9:691–704.

[109] Ahmed F, Bari ASMH, Gavrilova ML. Emotion Recognition From Body Movement. *IEEE Access*. 2020;8:11761–11781.

[110] Gumelar AB, Kurniawan A, Sooai AG, et al. Human Voice Emotion Identification Using Prosodic and Spectral Feature Extraction Based on Deep Neural Networks. 2019 IEEE 7th Int Conf Serious Games Appl Health SeGAH. 2019. p. 1–8.

[111] Fairbanks G, Hoaglin LW. An experimental study of the durational characteristics of the voice during the expression of emotion. *Speech Monogr*. 1941;8:85–90.

[112] Scherer KR. Expression of emotion in voice and music. *J Voice*. 1995;9:235–248.

[113] Levenson RW, Heider K, Ekman P. Emotion and Autonomic Nervous System Activity in the Minangkabau of West Sumatra. :17.

[114] Rabrait C. *Imagerie par résonance magnétique à haute résolution temporelle: Développement d'une méthode d'acquisition parallèle tridimensionnelle pour l'imagerie fonctionnelle cérébrale* [Internet] [phdthesis]. Université Paris Sud - Paris XI; 2007 [cited 2022 Aug 31]. Available from: <https://tel.archives-ouvertes.fr/tel-00204555>.

[115] Andreelli F, Mosbah H. *IRM fonctionnelle cérébrale : les principes*. Médecine Mal Métaboliques. 2014;8:13–19.

[116] Hoa Denis. *L'IRM: pas à pas*. Montpellier: Sauramps médical Campus medica; 2008.

[117] Teplan M. FUNDAMENTALS OF EEG MEASUREMENT. *Meas Sci Rev*. 2002;2:11.

[118] Stern JM. *Atlas of EEG Patterns*. Lippincott Williams & Wilkins; 2005.

- [119] Hot P, Delplanque S. *Electrophysiologie de la cognition*. Dunod; 2013.
- [120] Vion-Dury J, Balzani C, Micoulaud-Franchi J-A. *Neurophysiologie clinique en psychiatrie* [Internet]. Elsevier Masson SAS. 2015 [cited 2022 Aug 31]. Available from: <https://www.elsevier-masson.fr/neurophysiologie-clinique-en-psychiatrie-9782294740862.html>.
- [121] Hausser-Hauw C. *Manuel d'EEG de l'adulte* [Internet]. Elsevier Masson SAS. 2007 [cited 2022 Aug 31]. Available from: <https://www.elsevier-masson.fr/manuel-deeg-de-ladulte-9782294071454.html>.
- [122] Razani J, Davidson TM, Murphy C. Smell Disorders. In: Gershwin ME, Incaudo GA, editors. *Dis Sinuses Compr Textb Diagn Treat* [Internet]. Totowa, NJ: Humana Press; 1996 [cited 2022 Oct 25]. p. 425–450. Available from: https://doi.org/10.1007/978-1-4612-0225-7_22.
- [123] Jafek BW, Gordon AS, Moran DT, et al. Congenital anosmia. *Ear Nose Throat J*. 1990;69:331–337.
- [124] Vowles RH, Bleach NR, Rowe-Jones JM. Congenital anosmia. *Int J Pediatr Otorhinolaryngol*. 1997;41:207–214.
- [125] Mackay-Sim A, Johnston ANB, Owen C, et al. Olfactory Ability in the Healthy Population: Reassessing Presbyosmia. *Chem Senses*. 2006;31:763–771.
- [126] Oliva AD, Gupta R, Issa K, et al. Aging-related olfactory loss is associated with olfactory stem cell transcriptional alterations in humans. *J Clin Invest* [Internet]. 2022 [cited 2022 Oct 25];132. Available from: <https://www.jci.org/articles/view/155506>.
- [127] Goncalves S, Goldstein BJ. Pathophysiology of Olfactory Disorders and Potential Treatment Strategies. *Curr Otorhinolaryngol Rep*. 2016;4:115–121.
- [128] Zusho H. Posttraumatic Anosmia. *Arch Otolaryngol*. 1982;108:90–92.
- [129] Meng X, Deng Y, Dai Z, et al. COVID-19 and anosmia: A review based on up-to-date knowledge. *Am J Otolaryngol*. 2020;41:102581.
- [130] Gaines AD. Anosmia and hyposmia. *Ocean Publ*. 2010;31:185.
- [131] Hawkes C. Olfaction in Neurodegenerative Disorder. *Taste Smell*. 2006;63:133–151.
- [132] Joussain P, Bessy M, Faure F, et al. Application of the European Test of Olfactory Capabilities in patients with olfactory impairment. *Eur Arch Otorhinolaryngol*. 2016;273:381–390.
- [133] Klopfenstein T, Kadiane-Oussou NJ, Toko L, et al. Features of anosmia in COVID-19. *Médecine Mal Infect*. 2020;50:436–439.
- [134] Boesveldt S, Postma EM, Boak D, et al. Anosmia—A Clinical Review. *Chem Senses*. 2017;42:513–523.
- [135] Croy I, Nordin S, Hummel T. Olfactory Disorders and Quality of Life—An Updated Review. *Chem Senses*. 2014;39:185–194.
- [136] Kollndorfer K, Reichert J l., Brückler B, et al. Self-esteem as an important factor in quality of life and depressive symptoms in anosmia: A pilot study. *Clin Otolaryngol*. 2017;42:1229–1234.
- [137] Smeets MAM, Veldhuizen MG, Galle S, et al. Sense of smell disorder and health-related quality of life. *Rehabil Psychol*. 2009;54:404–412.
- [138] Patel HK. Biological Olfaction. In: Patel HK, editor. *Electron Nose Artif Olfaction Technol* [Internet]. New Delhi: Springer India; 2014 [cited 2022 Oct 25]. p. 67–84. Available from: https://doi.org/10.1007/978-81-322-1548-6_3.
- [139] Stevenson RJ. An Initial Evaluation of the Functions of Human Olfaction. *Chem Senses*. 2010;35:3–20.
- [140] Doty RL. *Handbook of Olfaction and Gustation*. John Wiley & Sons; 2015.
- [141] Richard Ortegón S, Fournel A, Carlos O, et al. And I'm feeling good: effect of emotional sweat and perfume on others' physiology, verbal responses, and creativity. *Chem*

Senses. 2022;47:bjac012.

[142] Gildersleeve KA, Haselton MG, Larson CM, et al. Body odor attractiveness as a cue of impending ovulation in women: Evidence from a study using hormone-confirmed ovulation. *Horm Behav.* 2012;61:157–166.

[143] Ferdenzi C, Richard Ortegón S, Delplanque S, et al. Interdisciplinary challenges for elucidating human olfactory attractiveness. *Philos Trans R Soc B Biol Sci.* 2020;375:20190268.

[144] de Groot JHB, Semin GR, Smeets MAM. On the Communicative Function of Body Odors: A Theoretical Integration and Review. *Perspect Psychol Sci.* 2017;12:306–324.

[145] Mori K, Yoshihara Y. Molecular recognition and olfactory processing in the mammalian olfactory system. *Prog Neurobiol.* 1995;45:585–619.

[146] Proust B. *Petite Géométrie des parfums.* Média Diffusion; 2009.

[147] Pelosi P. Odorant-Binding Proteins. *Crit Rev Biochem Mol Biol.* 1994;29:199–228.

[148] Bogert MT. Factors upon which Odor Appears to Depend. Factors Which Odor Appears Depend [Internet]. Columbia University Press; 1927 [cited 2022 Oct 25]. p. 9–13. Available from: <https://www.degruyter.com/document/doi/10.7312/boge93008-005/html>.

[149] Johnson BA, Leon M. Chemotopic odorant coding in a mammalian olfactory system. *J Comp Neurol.* 2007;503:1–34.

[150] Brookes JC, Horsfield AP, Stoneham AM. Odour character differences for enantiomers correlate with molecular flexibility. *J R Soc Interface.* 2009;6:75–86.

[151] Newman MP, Féron F, Mackay-Sim A. Growth factor regulation of neurogenesis in adult olfactory epithelium. *Neuroscience.* 2000;99:343–350.

[152] Purves D, Augustine GJ, Fitzpatrick D, et al. *Neurosciences.* De Boeck Supérieur; 2019.

[153] Breipohl W, Mackay-Sim A, Grandt D, et al. Neurogenesis in the Vertebrate Main Olfactory Epithelium. In: Breipohl W, Apfelbach R, editors. *Ontog Olfaction Princ Olfactory Matur Vertebr* [Internet]. Berlin, Heidelberg: Springer; 1986 [cited 2022 Oct 28]. p. 21–33. Available from: https://doi.org/10.1007/978-3-642-71576-1_2.

[154] Bushdid C, Magnasco MO, Vosshall LB, et al. Humans can Discriminate more than one Trillion Olfactory Stimuli. *Science.* 2014;343:1370–1372.

[155] Kleene SJ. The Electrochemical Basis of Odor Transduction in Vertebrate Olfactory Cilia. *Chem Senses.* 2008;33:839–859.

[156] Sézille C, Bensafi M. De la molécule au percept. *Molécule Au Percept.* 2013;24–26.

[157] Zatorre RJ, Jones-Gotman M. 13 - Functional Imaging of the Chemical Senses. In: Toga AW, Mazziotta JC, editors. *Brain Mapp Syst* [Internet]. San Diego: Academic Press; 2000 [cited 2022 Sep 1]. p. 403–424. Available from: <https://www.sciencedirect.com/science/article/pii/B9780126925456500155>.

[158] Gottfried JA. Central mechanisms of odour object perception. *Nat Rev Neurosci.* 2010;11:628–641.

[159] Holley A. Système olfactif et neurobiologie. *Terrain Anthropol Sci Hum.* 2006;107–122.

[160] Aggleton JP, Mishkin M. Chapter 12 - THE AMYGDALA: SENSORY GATEWAY TO THE EMOTIONS. In: Plutchik R, Kellerman H, editors. *Biol Found Emot* [Internet]. Academic Press; 1986 [cited 2022 Aug 11]. p. 281–299. Available from: <https://www.sciencedirect.com/science/article/pii/B9780125587037500188>.

[161] Yeshurun Y, Lapid H, Dudai Y, et al. The Privileged Brain Representation of First Olfactory Associations. *Curr Biol.* 2009;19:1869–1874.

[162] Soudry Y, Lemogne C, Malinvaud D, et al. Olfactory system and emotion: Common substrates. *Eur Ann Otorhinolaryngol Head Neck Dis.* 2011;128:18–23.

[163] Zhou W, Chen D. Sociochemosensory and Emotional Functions: Behavioral

- Evidence for Shared Mechanisms. *Psychol Sci.* 2009;20:1118–1124.
- [164] Roque LS, Kendrick KH, Norcliffe E, et al. Vision verbs dominate in conversation across cultures, but the ranking of non-visual verbs varies. *Cogn Linguist.* 2015;26:31–60.
- [165] Winter B. Taste and smell words form an affectively loaded and emotionally flexible part of the English lexicon. *Lang Cogn Neurosci.* 2016;31:975–988.
- [166] Dam-Jensen H, Zethsen KK. Pragmatic patterns and the lexical system—A reassessment of evaluation in language. *J Pragmat.* 2007;39:1608–1623.
- [167] Krifka M. A Note on an Asymmetry in the Hedonic Implicatures of Olfactory and Gustatory Terms. 2010;235–245.
- [168] Levinson SC, Majid A. Differential Ineffability and the Senses. *Mind Lang.* 2014;29:407–427.
- [169] Engen T. Remembering Odors and Their Names. *Am Sci.* 1987;75:497–503.
- [170] Dubois D. Categories as acts of meaning: The case of categories in Olfaction and audition. *Cogn Sci Q.* 2000;1:35–68.
- [171] Rouby C, Schaal B, Dubois D, et al. *Olfaction, Taste, and Cognition.* Cambridge University Press; 2002.
- [172] David S, Dubois D, Rouby C, et al. L’expression des odeurs en français : analyse lexicale et représentation cognitive. *Intellectica.* 1997;24:51–83.
- [173] Licon CC, Manesse C, Dantec M, et al. Pleasantness and trigeminal sensations as salient dimensions in organizing the semantic and physiological spaces of odors. *Sci Rep.* 2018;8:8444.
- [174] Zarzo M. Psychologic Dimensions in the Perception of Everyday Odors: Pleasantness and Edibility. *J Sens Stud.* 2008;23:354–376.
- [175] Royet J-P, Zald D, Versace R, et al. Emotional Responses to Pleasant and Unpleasant Olfactory, Visual, and Auditory Stimuli: a Positron Emission Tomography Study. *J Neurosci.* 2000;20:7752–7759.
- [176] Gottfried JA, Deichmann R, Winston JS, et al. Functional Heterogeneity in Human Olfactory Cortex: An Event-Related Functional Magnetic Resonance Imaging Study. *J Neurosci.* 2002;22:10819–10828.
- [177] Anderson AK, Christoff K, Stappen I, et al. Dissociated neural representations of intensity and valence in human olfaction. *Nat Neurosci.* 2003;6:196–202.
- [178] Royet J-P, Plailly J, Delon-Martin C, et al. fMRI of emotional responses to odors:: influence of hedonic valence and judgment, handedness, and gender. *NeuroImage.* 2003;20:713–728.
- [179] Kringelbach ML. The human orbitofrontal cortex: linking reward to hedonic experience. *Nat Rev Neurosci.* 2005;6:691–702.
- [180] O’Doherty J, Rolls ET, Francis S, et al. Sensory-specific satiety-related olfactory activation of the human orbitofrontal cortex. *NeuroReport.* 2000;11:893–897.
- [181] Rolls ET. Taste, olfactory, and food reward value processing in the brain. *Prog Neurobiol.* 2015;127–128:64–90.
- [182] Zald DH, Pardo JV. Emotion, olfaction, and the human amygdala: Amygdala activation during aversive olfactory stimulation. *Proc Natl Acad Sci.* 1997;94:4119–4124.
- [183] Bensafi M, Sobel N, Khan RM. Hedonic-Specific Activity in Piriform Cortex During Odor Imagery Mimics That During Odor Perception. *J Neurophysiol.* 2007;98:3254–3262.
- [184] Zelano C, Montag J, Johnson B, et al. Dissociated Representations of Irritation and Valence in Human Primary Olfactory Cortex. *J Neurophysiol.* 2007;97:1969–1976.
- [185] Gottfried JA, Dolan RJ. The Nose Smells What the Eye Sees: Crossmodal Visual Facilitation of Human Olfactory Perception. *Neuron.* 2003;39:375–386.
- [186] Howard JD, Plailly J, Grueschow M, et al. Odor quality coding and categorization in human posterior piriform cortex. *Nat Neurosci.* 2009;12:932–938.

- [187] Rolls ET, Kringelbach ML, De Araujo IET. Different representations of pleasant and unpleasant odours in the human brain. *Eur J Neurosci.* 2003;18:695–703.
- [188] Midroit M, Chalencon L, Renier N, et al. Neural processing of the reward value of pleasant odorants. *Curr Biol.* 2021;31:1592-1605.e9.
- [189] Bensafi M, Iannilli E, Poncelet J, et al. Dissociated Representations of Pleasant and Unpleasant Olfacto-Trigeminal Mixtures: An fMRI Study. *PLOS ONE.* 2012;7:e38358.
- [190] Fulbright RK, Skudlarski P, Lacadie CM, et al. Functional MR imaging of regional brain responses to pleasant and unpleasant odors. *Am J Neuroradiol.* 1998;19:1721–1726.
- [191] Winston JS, Gottfried JA, Kilner JM, et al. Integrated Neural Representations of Odor Intensity and Affective Valence in Human Amygdala. *J Neurosci.* 2005;25:8903–8907.
- [192] Masaoka Y, Koiwa N, Homma I. Inspiratory phase-locked alpha oscillation in human olfaction: source generators estimated by a dipole tracing method. *J Physiol.* 2005;566:979–997.
- [193] Alaoui-Ismaïli O, Vernet-Maury E, Dittmar A, et al. Odor Hedonics: Connection With Emotional Response Estimated by Autonomic Parameters. *Chem Senses.* 1997;22:237–248.
- [194] He W, Boesveldt S, de Graaf C, et al. Dynamics of autonomic nervous system responses and facial expressions to odors. *Front Psychol [Internet].* 2014 [cited 2021 Feb 16];5. Available from: <https://www.frontiersin.org/articles/10.3389/fpsyg.2014.00110/full>.
- [195] Bensafi M, Rouby C, Farget V, et al. Autonomic Nervous System Responses to Odours: the Role of Pleasantness and Arousal. *Chem Senses.* 2002;27:703–709.
- [196] Brauchli P, Regg PB, Etzweiler F, et al. Electrocortical and Autonomic Alteration by Administration of a Pleasant and an Unpleasant Odor. *Chem Senses.* 1995;20:505–515.
- [197] Alaoui-Ismaïli O, Robin O, Rada H, et al. Basic Emotions Evoked by Odorants: Comparison Between Autonomic Responses and Self-Evaluation. *Physiol Behav.* 1997;62:713–720.
- [198] Robin O, Alaoui-Ismaïli O, Dittmar A, et al. Emotional Responses Evoked by Dental Odors: An Evaluation from Autonomic Parameters. *J Dent Res.* 1998;77:1638–1646.
- [199] Brand G, Millot J-L, Biju C. Comparison between monorhinal and birhinal olfactory stimulations in bilateral electrodermal recordings. *Comptes Rendus Acadmie Sci - Ser III - Sci Vie.* 2000;323:959–965.
- [200] Glass ST, Lingg E, Heuberger E. Do ambient urban odors evoke basic emotions? *Front Psychol [Internet].* 2014 [cited 2022 Sep 1];5. Available from: <https://www.frontiersin.org/articles/10.3389/fpsyg.2014.00340>.
- [201] Delompr T, Leloup C, Salles C, et al. Got et olfaction : des rcepteurs au cerveau. *Mdecine Mal Mtaboliques.* 2022;16:491–501.
- [202] Bensafi M, Rouby C. Cerveau et odorat 1er dition: Comment (r)duquer son nez [Internet]. *Cerveau Odorat 1er dition.* EDP Sciences; 2020 [cited 2022 Oct 28]. Available from: <https://www.degruyter.com/document/doi/10.1051/978-2-7598-2509-7/html?lang=en>.
- [203] Mainland JD, Keller A, Li YR, et al. The Missense of Smell: Functional Variability in the Human Odorant Receptor Repertoire. *Nat Neurosci.* 2014;17:114–120.
- [204] Amoore JE, Venstrom D, Davis AR. Measurement of Specific Anosmia. *Percept Mot Skills.* 1968;26:143–164.
- [205] AMOORE JE. SPECIFIC ANOSMIA AND THE CONCEPT OF PRIMARY ODORS. *Chem Senses.* 1977;2:267–281.
- [206] Bonneau M. Use of entire males for pig meat in the European Union. *Meat Sci.* 1998;49:S257–S272.
- [207] Wysocki CJ, Dorries KM, Beauchamp GK. Ability to perceive androstenone can be

- acquired by ostensibly anosmic people. *Proc Natl Acad Sci.* 1989;86:7976–7978.
- [208] Bonneau M, Chevillon P. Acceptability of entire male pork with various levels of androstenone and skatole by consumers according to their sensitivity to androstenone. *Meat Sci.* 2012;90:330–337.
- [209] Lunde K, Egelanddal B, Skuterud E, et al. Genetic Variation of an Odorant Receptor OR7D4 and Sensory Perception of Cooked Meat Containing Androstenone. *PLOS ONE.* 2012;7:e35259.
- [210] Keller A, Zhuang H, Chi Q, et al. Genetic variation in a human odorant receptor alters odour perception. *Nature.* 2007;449:468–472.
- [211] Sorokowski P, Karwowski M, Misiak M, et al. Sex Differences in Human Olfaction: A Meta-Analysis. *Front Psychol* [Internet]. 2019 [cited 2022 Apr 15];10. Available from: <https://www.frontiersin.org/article/10.3389/fpsyg.2019.00242>.
- [212] Allen C, Cobey KD, Havlíček J, et al. The impact of artificial fragrances on the assessment of mate quality cues in body odor. *Evol Hum Behav.* 2016;37:481–489.
- [213] Mair RG, Bouffard JA, Engen T, et al. Olfactory sensitivity during the menstrual cycle. *Sens Processes.* 1978;2:90–98.
- [214] Hummel T, Gollisch T, Wildt R, et al. Changes in olfactory perception during the menstrual cycle. *Experientia.* 1991;47:712–715.
- [215] Doty RL, Cameron EL. Sex differences and reproductive hormone influences on human odor perception. *Physiol Behav.* 2009;97:213–228.
- [216] Rouby C, Pouliot S, Bensafi M. Odor hedonics and their modulators. *Food Qual Prefer.* 2009;20:545–549.
- [217] Pouliot S, Bourgeat F, Barkat S, et al. Increase in Anhedonia Level in Menopausal Women is Accompanied by a Shift in Olfactory Function. *Chemosens Percept.* 2008;1:43–47.
- [218] Sorokowska A, Sorokowski P, Hummel T, et al. Olfaction and Environment: Tsimane’ of Bolivian Rainforest Have Lower Threshold of Odor Detection Than Industrialized German People. *PLOS ONE.* 2013;8:e69203.
- [219] Bensafi M, Rouby C. Chapitre 2. Perception olfactive : entre universalité et diversité. Chapitre 2 Percept Olfactive Entre Universalité Divers [Internet]. EDP Sciences; 2021 [cited 2022 Aug 17]. p. 33–56. Available from: <https://www.degruyter.com/document/doi/10.1051/978-2-7598-2639-1.c004/html>.
- [220] Ferdenzi C, Roberts SC, Schirmer A, et al. Variability of Affective Responses to Odors: Culture, Gender, and Olfactory Knowledge. *Chem Senses.* 2013;38:175–186.
- [221] Schaal B. Variability and invariants in early odour preferences : comparative data from children belonging to three cultures. *Chem Senses.* 1997;22:212.
- [222] Stone H. Influence of temperature on olfactory sensitivity. *J Appl Physiol.* 1963;18:746–751.
- [223] Kuehn M, Welsch H, Zahnert T, et al. Changes of pressure and humidity affect olfactory function. *Eur Arch Otorhinolaryngol.* 2008;265:299–302.
- [224] Sorokowska A, Sorokowski P, Frackowiak T. Determinants of human olfactory performance: A cross-cultural study. *Sci Total Environ.* 2015;506–507:196–200.
- [225] Ajmani GS, Suh HH, Pinto JM. Effects of Ambient Air Pollution Exposure on Olfaction: A Review. *Environ Health Perspect.* 2016;124:1683–1693.
- [226] Katotomichelakis M, Balatsouras D, Tripsianis G, et al. The effect of smoking on the olfactory function. *Rhinology.* 2008;45:273–280.
- [227] Vent J, Robinson AM, Gentry-Nielsen MJ, et al. Pathology of the Olfactory Epithelium: Smoking and Ethanol Exposure. *The Laryngoscope.* 2004;114:1383–1388.
- [228] Ajmani GS, Suh HH, Wroblewski KE, et al. Smoking and Olfactory Dysfunction: a Systematic Literature Review and Meta-Analysis. *The Laryngoscope.* 2017;127:1753–1761.

- [229] Wongrakpanich S, Petchlorlian A, Rosenzweig A. Sensorineural Organs Dysfunction and Cognitive Decline: A Review. *Aging Dis.* 2016;7:763.
- [230] Doty R, Kamath V. The influences of age on olfaction: a review. *Front Psychol* [Internet]. 2014 [cited 2022 Apr 15];5. Available from: <https://www.frontiersin.org/article/10.3389/fpsyg.2014.00020>.
- [231] Sugiura TA Junko Mori, Yoshiaki Nakai, Midori. An Epidemiological Study of Postviral Olfactory Disorder. *Acta Otolaryngol (Stockh)*. 1998;118:191–196.
- [232] Mehraeen E, Behnezhad F, Salehi MA, et al. Olfactory and gustatory dysfunctions due to the coronavirus disease (COVID-19): a review of current evidence. *Eur Arch Otorhinolaryngol*. 2021;278:307–312.
- [233] ter LH, Renkawek K, van WF. The olfactory bulb in Alzheimer disease: a morphologic study of neuron loss, tangles, and senile plaques in relation to olfaction. *Alzheimer Dis Assoc Disord*. 1994;8:38–48.
- [234] Doty RL. Olfactory dysfunction in Parkinson disease. *Nat Rev Neurol*. 2012;8:329–339.
- [235] Chen B, Klarmann R, Israel M, et al. Difference of olfactory deficit in patients with acute episode of schizophrenia and major depressive episode. *Schizophr Res*. 2019;212:99–106.
- [236] Croy I, Hummel T. Olfaction as a marker for depression. *J Neurol*. 2017;264:631–638.
- [237] Joussain P, Thevenet M, Rouby C, et al. Effect of Aging on Hedonic Appreciation of Pleasant and Unpleasant Odors. *PLOS ONE*. 2013;8:e61376.
- [238] Elsner R. Environment and medication use influence olfactory abilities of older adults. *J Nutr Health Aging*. 2001;5:5–10.
- [239] Douglass R, Heckman G. Drug-related taste disturbance: A contributing factor in geriatric syndromes. *Can Fam Physician*. 2010;56:1142–1147.
- [240] Thiermann F, Buchbauer G. Influence of drugs on smelling capacity. A review. *Flavour Fragr J*. 2017;32:147–158.
- [241] Pollatos O, Kopietz R, Linn J, et al. Emotional Stimulation Alters Olfactory Sensitivity and Odor Judgment. *Chem Senses*. 2007;32:583–589.
- [242] Hoenen M, Wolf OT, Pause BM. The Impact of Stress on Odor Perception. *Perception*. 2017;46:366–376.
- [243] Cain WS, Johnson F. Lability of Odor Pleasantness: Influence of Mere Exposure. *Perception*. 1978;7:459–465.
- [244] Rolls BJ. Sensory-specific satiety. *Nutr Rev USA* [Internet]. 1986 [cited 2022 Oct 28]; Available from: https://scholar.google.com/scholar_lookup?title=Sensory-specific+satiety&author=Rolls%2C+B.J.&publication_year=1986.
- [245] Fernandez P, Bensafi M, Rouby C, et al. Does olfactory specific satiety take place in a natural setting? *Appetite*. 2013;60:1–4.
- [246] Stafford LD, Welbeck K. High Hunger State Increases Olfactory Sensitivity to Neutral but Not Food Odors. *Chem Senses*. 2011;36:189–198.
- [247] Delplanque S, Grandjean D, Chrea C, et al. Sequential unfolding of novelty and pleasantness appraisals of odors: Evidence from facial electromyography and autonomic reactions. *Emotion*. 2009;9:316–328.
- [248] Watanuki S, Kim Y-K. Physiological Responses Induced by Pleasant Stimuli. *J Physiol Anthropol Appl Human Sci*. 2005;24:135–138.
- [249] Møller P, Dijksterhuis G. Differential human electrodermal responses to odours. *Neurosci Lett*. 2003;346:129–132.
- [250] Schneider F, Grodd W, Weiss U, et al. Functional MRI reveals left amygdala activation during emotion. *Psychiatry Res Neuroimaging*. 1997;76:75–82.

- [251] Bensafi M, Rouby C, Farget V, et al. Psychophysiological correlates of affects in human olfaction. *Neurophysiol Clin Neurophysiol*. 2002;32:326–332.
- [252] Kroupi E, Vesin J, Ebrahimi T. Subject-Independent Odor Pleasantness Classification Using Brain and Peripheral Signals. *IEEE Trans Affect Comput*. 2016;7:422–434.
- [253] Kragel PA, LaBar KS. Multivariate Pattern Classification Reveals Autonomic and Experiential Representations of Discrete Emotions. *Emot Wash DC*. 2013;13:681–690.
- [254] Keller A, Gerkin RC, Guan Y, et al. Predicting human olfactory perception from chemical features of odor molecules. *Science*. 2017;355:820–826.
- [255] Khan RM, Luk C-H, Flinker A, et al. Predicting Odor Pleasantness from Odorant Structure: Pleasantness as a Reflection of the Physical World. *J Neurosci*. 2007;27:10015–10023.
- [256] Becerra MA, Londoño-Delgado E, Pelaez-Becerra SM, et al. Odor Pleasantness Classification from Electroencephalographic Signals and Emotional States. In: Serrano C, JE, Martínez-Santos JC, editors. *Adv Comput*. Cham: Springer International Publishing; 2018. p. 128–138.
- [257] Desbois D. L'intelligence artificielle. De quoi s'agit-il vraiment ? *Creis-Terminal*; 2021.
- [258] Berkhin P. A Survey of Clustering Data Mining Techniques. In: Kogan J, Nicholas C, Teboulle M, editors. *Group Multidimens Data Recent Adv Clust* [Internet]. Berlin, Heidelberg: Springer; 2006 [cited 2022 Sep 1]. p. 25–71. Available from: https://doi.org/10.1007/3-540-28349-8_2.
- [259] Aggarwal CC. An Introduction to Data Mining. In: Aggarwal CC, editor. *Data Min Textb* [Internet]. Cham: Springer International Publishing; 2015 [cited 2022 Aug 24]. p. 1–26. Available from: https://doi.org/10.1007/978-3-319-14142-8_1.
- [260] Fayyad U, Piatetsky-Shapiro G, Smyth P. From Data Mining to Knowledge Discovery in Databases. *AI Mag*. 1996;17:37–37.
- [261] Guyon I, Elisseeff A. An Introduction of Variable and Feature Selection. *J Mach Learn Res Spec Issue Var Feature Sel*. 2003;3:1157–1182.
- [262] Novak PK, Lavrač N, Webb GI. Supervised Descriptive Rule Discovery: A Unifying Survey of Contrast Set, Emerging Pattern and Subgroup Mining. *J Mach Learn Res*. 2009;10:377–403.
- [263] Zaki MJ, Jr WM. *Data Mining and Machine Learning: Fundamental Concepts and Algorithms*. Cambridge University Press; 2020.
- [264] Wrobel S. An algorithm for multi-relational discovery of subgroups. In: Komorowski J, Zytkow J, editors. *Princ Data Min Knowl Discov*. Berlin, Heidelberg: Springer; 1997. p. 78–87.
- [265] Atzmueller M. Subgroup discovery. *WIREs Data Min Knowl Discov*. 2015;5:35–49.
- [266] Herrera F, Carmona CJ, González P, et al. An overview on subgroup discovery: foundations and applications. *Knowl Inf Syst*. 2011;29:495–525.
- [267] Leman D, Feelders A, Knobbe A. Exceptional Model Mining. In: Daelemans W, Goethals B, Morik K, editors. *Mach Learn Knowl Discov Databases*. Berlin, Heidelberg: Springer; 2008. p. 1–16.
- [268] Lavrac N, Kavsek B, Flach P, et al. Subgroup Discovery with CN2-SD. :36.
- [269] Sasaki Y. The truth of the F-measure. *Teach Tutor Mater*. 1:1–5.
- [270] Niwattanakul S, Singthongchai J, Naenudorn E, et al. Using of Jaccard Coefficient for Keywords Similarity. 2013.
- [271] Belfodil A, Belfodil A, Kaytoute M. Anytime Subgroup Discovery in Numerical Domains with Guarantees. In: Berlingerio M, Bonchi F, Gärtner T, et al., editors. *Mach Learn Knowl Discov Databases*. Cham: Springer International Publishing; 2019. p. 500–

516.

[272] McNicholas PD, Murphy TB, O'Regan M. Standardising the lift of an association rule. *Comput Stat Data Anal.* 2008;52:4712–4721.

[273] Boley M, Lucchese C, Paurat D, et al. Direct local pattern sampling by efficient two-step random procedures. *Proc 17th ACM SIGKDD Int Conf Knowl Discov Data Min [Internet]*. New York, NY, USA: Association for Computing Machinery; 2011 [cited 2022 Sep 1]. p. 582–590. Available from: <https://doi.org/10.1145/2020408.2020500>.

[274] Boley M, Moens S, Gärtner T. Linear space direct pattern sampling using coupling from the past. *Proc 18th ACM SIGKDD Int Conf Knowl Discov Data Min [Internet]*. New York, NY, USA: Association for Computing Machinery; 2012 [cited 2022 Sep 1]. p. 69–77. Available from: <https://doi.org/10.1145/2339530.2339545>.

[275] Moens S, Boley M. Instant Exceptional Model Mining Using Weighted Controlled Pattern Sampling. In: Blockeel H, van Leeuwen M, Vinciotti V, editors. *Adv Intell Data Anal XIII*. Cham: Springer International Publishing; 2014. p. 203–214.

[276] del Jesus MJ, Gonzalez P, Herrera F, et al. Evolutionary Fuzzy Rule Induction Process for Subgroup Discovery: A Case Study in Marketing. *IEEE Trans Fuzzy Syst.* 2007;15:578–592.

[277] Rodríguez D, Ruiz R, Riquelme JC, et al. Searching for rules to detect defective modules: A subgroup discovery approach. *Inf Sci.* 2012;191:14–30.

[278] Bosc G, Boulicaut J-F, Raïssi C, et al. Anytime discovery of a diverse set of patterns with Monte Carlo tree search. *Data Min Knowl Discov.* 2018;32:604–650.

[279] Bosc G, Golebiowski J, Bensafi M, et al. Local Subgroup Discovery for Eliciting and Understanding New Structure-Odor Relationships. In: Calders T, Ceci M, Malerba D, editors. *Discov Sci*. Cham: Springer International Publishing; 2016. p. 19–34.

[280] Bosc G, Plantevit M, Boulicaut J-F, et al. h(odor): Interactive Discovery of Hypotheses on the Structure-Odor Relationship in Neuroscience. In: Berendt B, Bringmann B, Fromont É, et al., editors. *Mach Learn Knowl Discov Databases*. Cham: Springer International Publishing; 2016. p. 17–21.

[281] Licon CC, Bosc G, Sabri M, et al. Chemical features mining provides new descriptive structure-odor relationships. *PLOS Comput Biol.* 2019;15:e1006945.

[282] Moranges M, Plantevit M, Fournel A, et al. Exceptional Attributed Subgraph Mining to Understand the Olfactory Percept. In: Soldatova L, Vanschoren J, Papadopoulos G, et al., editors. *Discov Sci*. Cham: Springer International Publishing; 2018. p. 276–291.

[283] de Wijk RA, Noldus LPJJ. Using implicit rather than explicit measures of emotions. *Food Qual Prefer.* 2021;92:104125.

[284] Croy I, Drechsler E, Hamilton P, et al. Olfactory modulation of affective touch processing — A neurophysiological investigation. *NeuroImage.* 2016;135:135–141.

[285] Seet MS, Amin MdR, Abbasi NI, et al. Olfactory-induced Positive Affect and Autonomic Response as a Function of Hedonic and Intensity Attributes of Fragrances. 2020 42nd Annu Int Conf IEEE Eng Med Biol Soc EMBC. 2020. p. 3170–3173.

[286] Kadohisa M. Effects of odor on emotion, with implications. *Front Syst Neurosci [Internet]*. 2013 [cited 2022 Nov 2];7. Available from: <https://www.frontiersin.org/articles/10.3389/fnsys.2013.00066>.

[287] Delplanque S, Coppin G, Sander D. Odor and Emotion. In: Buettner A, editor. *Springer Handb Odor [Internet]*. Cham: Springer International Publishing; 2017 [cited 2022 Nov 2]. p. 101–102. Available from: https://doi.org/10.1007/978-3-319-26932-0_40.

[288] Pichon AM, Coppin G, Cayeux I, et al. Sensitivity of Physiological Emotional Measures to Odors Depends on the Product and the Pleasantness Ranges Used. *Front Psychol [Internet]*. 2015 [cited 2022 Nov 2];6. Available from: <https://www.frontiersin.org/articles/10.3389/fpsyg.2015.01821>.

- [289] Herz RS. Aromatherapy Facts and Fictions: A Scientific Analysis of Olfactory Effects on Mood, Physiology and Behavior. *Int J Neurosci*. 2009;119:263–290.
- [290] Herz RS, Eliassen J, Beland S, et al. Neuroimaging evidence for the emotional potency of odor-evoked memory. *Neuropsychologia*. 2004;42:371–378.
- [291] Jo H-G, Wudarczyk O, Leclerc M, et al. Effect of odor pleasantness on heat-induced pain: An fMRI study. *Brain Imaging Behav*. 2021;15:1300–1312.
- [292] Robin O, Alaoui-Ismaïli O, Dittmar A, et al. Basic Emotions Evoked by Eugenol Odor Differ According to the Dental Experience. A Neurovegetative Analysis. *Chem Senses*. 1999;24:327–335.
- [293] Hou H-R, Zhang X-N, Meng Q-H. Odor-induced emotion recognition based on average frequency band division of EEG signals. *J Neurosci Methods*. 2020;334:108599.
- [294] Mathonat R, Nurbakova D, Boulicaut J-F, et al. Anytime mining of sequential discriminative patterns in labeled sequences. *Knowl Inf Syst*. 2021;63:439–476.
- [295] Bendimerad A, Plantevit M, Robardet C. Mining exceptional closed patterns in attributed graphs. *Knowl Inf Syst*. 2018;56:1–25.
- [296] Djordjevic J, Lundstrom JN, Clément F, et al. A Rose by Any Other Name: Would it Smell as Sweet? *J Neurophysiol*. 2008;99:386–393.
- [297] Kato M, Okumura T, Tsubo Y, et al. Spatiotemporal dynamics of odor representations in the human brain revealed by EEG decoding. *Proc Natl Acad Sci*. 2022;119:e2114966119.
- [298] Bensafi M, Porter J, Pouliot S, et al. Olfactomotor activity during imagery mimics that during perception. *Nat Neurosci*. 2003;6:1142–1144.
- [299] Johnson BN, Mainland JD, Sobel N. Rapid Olfactory Processing Implicates Subcortical Control of an Olfactomotor System. *J Neurophysiol*. 2003;90:1084–1094.
- [300] Bensafi M, Rouby C, Farget V, et al. Perceptual, affective, and cognitive judgments of odors: Pleasantness and handedness effects. *Brain Cogn*. 2003;51:270–275.
- [301] Boesveldt S, Frasnelli J, Gordon AR, et al. The fish is bad: Negative food odors elicit faster and more accurate reactions than other odors. *Biol Psychol*. 2010;84:313–317.
- [302] Collet C, Vernet-Maury E, Delhomme G, et al. Autonomic nervous system response patterns specificity to basic emotions. *J Auton Nerv Syst*. 1997;62:45–57.
- [303] Li D, Jia J, Wang X. Unpleasant Food Odors Modulate the Processing of Facial Expressions: An Event-Related Potential Study. *Front Neurosci* [Internet]. 2020 [cited 2022 Nov 2];14. Available from: <https://www.frontiersin.org/articles/10.3389/fnins.2020.00686>.
- [304] Kline JP, Blackhart GC, Woodward KM, et al. Anterior electroencephalographic asymmetry changes in elderly women in response to a pleasant and an unpleasant odor. *Biol Psychol*. 2000;52:241–250.
- [305] Schneider F, Weiss U, Kessler C, et al. Subcortical correlates of differential classical conditioning of aversive emotional reactions in social phobia. *Biol Psychiatry*. 1999;45:863–871.
- [306] Kroupi E, Yazdani A, Vesin J-M, et al. EEG Correlates of Pleasant and Unpleasant Odor Perception. *ACM Trans Multimed Comput Commun Appl*. 2014;11:13:1-13:17.
- [307] Kim J, André E. Emotion recognition based on physiological changes in music listening. *IEEE Trans Pattern Anal Mach Intell*. 2008;30:2067–2083.
- [308] Neumann SA, Waldstein SR. Similar patterns of cardiovascular response during emotional activation as a function of affective valence and arousal and gender. *J Psychosom Res*. 2001;50:245–253.
- [309] Palomba D, Sarlo M, Angrilli A, et al. Cardiac responses associated with affective processing of unpleasant film stimuli. *Int J Psychophysiol*. 2000;36:45–57.
- [310] Vianna EPM, Tranel D. Gastric myoelectrical activity as an index of emotional arousal. *Int J Psychophysiol*. 2006;61:70–76.

- [311] de Wijk RA, Kooijman V, Verhoeven RHG, et al. Autonomic nervous system responses on and facial expressions to the sight, smell, and taste of liked and disliked foods. *Food Qual Prefer.* 2012;26:196–203.
- [312] Vrana SR. The psychophysiology of disgust: Differentiating negative emotional contexts with facial EMG. *Psychophysiology.* 1993;30:279–286.
- [313] Ritz T, Thöns M, Fahrenkrug S, et al. Airways, respiration, and respiratory sinus arrhythmia during picture viewing. *Psychophysiology.* 2005;42:568–578.
- [314] Ekman paul, Levenson RW. Autonomic nervous system activity distinguishes between emotions. *Science. Science.* 1983;221:1208–1210.
- [315] Hess U, Kappas A, McHugo GJ, et al. The facilitative effect of facial expression on the self-generation of emotion. *Int J Psychophysiol.* 1992;12:251–265.
- [316] Van Diest I, Proot P, Van De Woestijne KP, et al. Critical Conditions for Hyperventilation Responses: The Role of Autonomic Response Propositions During Emotional Imagery. *Behav Modif.* 2001;25:621–639.
- [317] Shiota M, Neufeld S, Yeung W, et al. Feeling Good: Autonomic Nervous System Responding in Five Positive Emotions. *Emot Wash DC.* 2011;11:1368–1378.
- [318] Haag A, Goronzy S, Schaich P, et al. Emotion Recognition Using Bio-sensors: First Steps towards an Automatic System. In: André E, Dybkjær L, Minker W, et al., editors. *Affect Dialogue Syst.* Berlin, Heidelberg: Springer; 2004. p. 36–48.
- [319] Lisetti CL, Nasoz F. Using Noninvasive Wearable Computers to Recognize Human Emotions from Physiological Signals. *EURASIP J Adv Signal Process.* 2004;2004:929414.
- [320] Yoo S, Lee C, Park Y, et al. Neural Network Based Emotion Estimation Using Heart Rate Variability and Skin Resistance. *Lect. Notes Comput. Sci.* 2005. p. 824.
- [321] Choi A, Woo W, Lab GU -vr, et al. W.: Physiological sensing and feature extraction for emotion recognition by exploiting acupuncture spots. *Acii 2005 Lncs.* Springer; 2005. p. 590–597.
- [322] Katsis CD, Katertsidis N, Ganiatsas G, et al. Toward Emotion Recognition in Car-Racing Drivers: A Biosignal Processing Approach. *IEEE Trans Syst Man Cybern - Part Syst Hum.* 2008;38:502–512.
- [323] Stephens CL, Christie IC, Friedman BH. Autonomic specificity of basic emotions: Evidence from pattern classification and cluster analysis. *Biol Psychol.* 2010;84:463–473.
- [324] Kroupi E, Sopic D, Ebrahimi T. Non-linear EEG features for odor pleasantness recognition. 2014 Sixth Int Workshop Qual Multimed Exp QoMEX. 2014. p. 147–152.
- [325] Yazdani A, Kroupi E, Vesin J-M, et al. Electroencephalogram alterations during perception of pleasant and unpleasant odors. 2012 Fourth Int Workshop Qual Multimed Exp. 2012. p. 272–277.
- [326] Liu C, Conn K, Sarkar N, et al. Physiology-based affect recognition for computer-assisted intervention of children with Autism Spectrum Disorder. *Int J Hum-Comput Stud.* 2008;66:662–677.
- [327] Zhai J, Barreto A. Stress Detection in Computer Users Based on Digital Signal Processing of Noninvasive Physiological Variables. 2006 Int Conf IEEE Eng Med Biol Soc. 2006. p. 1355–1358.
- [328] Frantzidis CA, Bratsas C, Klados MA, et al. On the Classification of Emotional Biosignals Evoked While Viewing Affective Pictures: An Integrated Data-Mining-Based Approach for Healthcare Applications. *IEEE Trans Inf Technol Biomed.* 2010;14:309–318.
- [329] Healey JA, Picard RW. Detecting stress during real-world driving tasks using physiological sensors. *IEEE Trans Intell Transp Syst.* 2005;6:156–166.
- [330] Shapley LS, Roth AE. *The Shapley value: essays in honor of Lloyd S. Shapley.* Cambridge University Press; 1988.
- [331] Di Flumeri G, Aricò P, Borghini G, et al. EEG-based Approach-Withdrawal index

- for the pleasantness evaluation during taste experience in realistic settings. 2017 39th Annu Int Conf IEEE Eng Med Biol Soc EMBC. 2017. p. 3228–3231.
- [332] Rogers LJ, Vallortigara G, Andrew RJ. *Divided Brains: The Biology and Behaviour of Brain Asymmetries*. Cambridge University Press; 2013.
- [333] Reimchen TE, Spoljaric MA. Right paw foraging bias in wild black bear (*Ursus americanus kermodei*). *Laterality*. 2011;16:471–478.
- [334] Hopkins WD, Bales SA, Bennett AJ. Heritability of Hand Preference in Chimpanzees (PAN). *Int J Neurosci*. 1994;74:17–26.
- [335] McGrew WC, Marchant LF. Laterality of hand use pays off in foraging success for wild chimpanzees. *Primates*. 1999;40:509–513.
- [336] Rodriguez A, Kaakinen M, Moilanen I, et al. Mixed-Handedness Is Linked to Mental Health Problems in Children and Adolescents. *Pediatrics*. 2010;125:e340–e348.
- [337] Weisberg J, McCullough S, Emmorey K. Simultaneous perception of a spoken and a signed language: The brain basis of ASL-English code-blends. *Brain Lang*. 2015;147:96–106.
- [338] Schnell AK, Hanlon RT, Benkada A, et al. Lateralization of Eye Use in Cuttlefish: Opposite Direction for Anti-Predatory and Predatory Behaviors. *Front Physiol* [Internet]. 2016 [cited 2022 Nov 15];0. Available from: <https://www.frontiersin.org/articles/10.3389/fphys.2016.00620/full>.
- [339] Joussain P, Ferdenzi C, Djordjevic J, et al. Relationship Between Psychophysiological Responses to Aversive Odors and Nutritional Status During Normal Aging. *Chem Senses*. 2017;42:465–472.
- [340] Pellegrino R, Sinding C, de Wijk RA, et al. Habituation and adaptation to odors in humans. *Physiol Behav*. 2017;177:13–19.
- [341] Profiling of Olfactory Receptor Gene Expression in Whole Human Olfactory Mucosa | PLOS ONE [Internet]. [cited 2022 Oct 31]. Available from: <https://journals.plos.org/plosone/article?id=10.1371/journal.pone.0096333>.
- [342] Abreu R, Leal A, Figueiredo P. EEG-Informed fMRI: A Review of Data Analysis Methods. *Front Hum Neurosci* [Internet]. 2018 [cited 2022 Oct 31];12. Available from: <https://www.frontiersin.org/articles/10.3389/fnhum.2018.00029>.
- [343] Ullsperger M, Debener S. *Simultaneous EEG and fMRI: Recording, Analysis, and Application*. Oxford University Press; 2010.
- [344] Williams LM, Phillips ML, Brammer MJ, et al. Arousal Dissociates Amygdala and Hippocampal Fear Responses: Evidence from Simultaneous fMRI and Skin Conductance Recording. *NeuroImage*. 2001;14:1070–1079.
- [345] Williams LM, Das P, Liddell B, et al. BOLD, sweat and fears: fMRI and skin conductance distinguish facial fear signals. *NeuroReport*. 2005;16:49–52.
- [346] Dangelmaier M, Blach R. Odor in Immersive Environments. In: Buettner A, editor. *Springer Handb Odor* [Internet]. Cham: Springer International Publishing; 2017 [cited 2022 Nov 2]. p. 139–140. Available from: https://doi.org/10.1007/978-3-319-26932-0_55.
- [347] Griffiths A. *Shivers Down Your Spine: Cinema, Museums, and the Immersive View*. Columbia University Press; 2008.
- [348] Muffat S, Garillon J-L. L’aromatique émotionnelle : une nouvelle dimension de la fonction olfactive. *Hegel*. 2014;3:287–295.

ANNEXES

Annexe 1 : « Étude de la variabilité de perception des odeurs à partir de descriptions subjectives » (soumis)

Language descriptions of smells are richer than thought

Mantel, M.^{1*}, Moranges, M.^{1,2}, Fournel, A.¹, Oelschlägel, A.³, Carro, J.¹, Hummel, T.³, Plantevit, M.², Roy, J.M.⁴, Bensafi, M.¹

¹ Lyon Neuroscience Research Center,
University of Lyon, CNRS UMR5292, INSERM, F-69366, France

² Laboratoire d'Informatique en Image et Systèmes d'information
CNRS UMR 5205, Lyon, France

³ Smell and Taste Clinic, Department of Otorhinolaryngology,
University of Dresden Medical School, Dresden, Germany

⁴ École Normale Supérieure de Lyon, Lyon, France

* Corresponding author: marylou.mantel@ens-lyon.fr

Abstract

Progress in the comprehension of smell perception requires a better account of its subjectivity, long neglected in contemporary cognitive science. Here, we aimed to investigate this issue with a multidisciplinary approach combining philosophy, sensory neuroscience, linguistics and artificial intelligence. We first developed an experimental protocol to collect subjective olfactory experience with less experimental biases, in a sample of French and German participants. In a second step, we developed a descriptive model including six main categories: source, impact, use, memory, qualitative description and difficulty. Although pleasantness and references to food were the most prevalent elements, categories were not equally used across odorants and individuals. Finally, using data mining algorithms, we showed that salient descriptors in discourse differed according to gender, language and exposure to odors. These findings show that olfactory experience is richer and more diverse than previously thought. They also stress the overall importance of subjective methodologies to understand fine-grained relationships between perception, language and the brain.

Keywords:

Olfaction - Perception - Interindividual Differences – Neurophenomenology – Data mining

1. Introduction

Perception is an essential function that allows us to detect and act upon relevant stimuli. Although vision and audition have been extensively studied in humans, the comprehension of smell is still lacking as it has been long considered as a minor sense, a remaining of evolution progressively replaced as the frontal regions developed (McGann, 2017). In fact, humans still have a keen sense of smell. It is especially important in food detection and assessment, hazard identification and in overall quality of life (Stevenson, 2010).

Beyond the dated idea of microsmatic humans, the growing interest in smell is confronted to the complexity of doing experimental research with odors. Indeed, olfactory scientists face three main issues. The first one is that the stimulus space is extremely large, with a variety of volatile compounds that can evoke a smell perception. Latest estimations suggest that humans can discriminate between more than a trillion of odorants (Buschdid et al., 2014). Chemists have long tried to identify the molecular features that are responsible for certain perceptions, but there is actually no simple association between the physicochemical structure of a molecule and the perception it evokes, a problem that has been conceptualized as the stimulus-percept issue (Auffarth, 2013; Keller et al., 2017; Khan et al., 2007; Licon et al., 2019). Secondly, humans are not equally equipped to detect odorants: around 400 types of receptors are present in the olfactory epithelium, but their nature depends on which genes are expressed by the individuals (Mainland et al., 2014). This has an impact on the perception of certain odorants, like beta-ionone (Jaeger et al., 2013) or androstenone (Keller et al., 2007), for which detection is determined by the expression of a specific allele combination, a phenomenon called specific anosmia.

The third and main barrier for olfactory researchers is the perception per se. Odors are known to be especially hard to talk about compared to colors or sounds, and smell is even considered as the “muted sense” for poetess Diane Ackerman. Indeed, odor identification performances are notoriously poor in humans (Cain, 1979). However, this type of task is actually not aimed to reflect the actual capacity to talk about odors, but rather the ability to associate a percept with a ‘veridical label’ assigned by the experimenter (Dubois & Rouby, 2002). In fact, when people describe an odor, they predominantly use emotional or source-based terms such as “flower” or “fruit”, as the vocabulary for odors is extremely limited in most Western languages. Therefore, olfactory perception is usually measured with quantitative ratings: how intense the odor is, how familiar, how pleasant and so on (Keller & Vosshall, 2004). When a qualitative description is needed, lists of semantic descriptors are often presented to the participants to choose from. Such descriptors are developed by experts in fields related to olfaction like perfumery or oenology, and although they allow for a standardized measure, they can be quite redundant and vague for novices (Dravnieks, 1985). Free description can also be used, but it is less common because of the variability in answers and the difficulties for individuals in verbalizing their experience of odors (Cain, 1979).

Ratings and forced-choice descriptors are very useful to get a standardized, quantifiable measurement of olfactory perception, but they are limited and leave out a large part of what constitutes the subjective olfactory experience. Indeed, they orient the participants towards very specific aspects of perception, like intensity or familiarity, which may not correspond to what they would actually pay attention to if not biased. Also, individuals often use their own words to describe odors, whether it is with souvenirs, metaphors or sensory descriptors, beyond the calibrated terms offered to them. These issues are problematic if olfactory researchers want to have a reliable measure of perception that include all its features, without the bias brought by the experimenter. The main aim of the present study was therefore to propose a new way to model olfactory perception by combining approaches from several areas of research. With the conceptual insights of philosophy of mind, we combined methods aimed specifically at collecting subjective accounts of phenomena with techniques from psychology and sensory neuroscience to construct the protocol, and discourse analysis for the exploitation of the linguistic data. In addition, artificial intelligence provided us with data mining tools for the fine-grained investigation of the verbalizations.

The stakes of such multidimensional approach are twofold. On the one hand, it will make it possible to revisit the hypothesis according to which human beings have a weak capacity for olfactory description. On the other hand, it will provide new scientific knowledge on the perceptive dimensions of smell, thus enabling the relationship between stimulus, biology and behavior to be studied in the least biased way possible.

In the literature, several models have been proposed to organize the structure of olfactory perception. Most dimensional reduction methods show the high importance of hedonics in the odor space. Pleasantness is indeed an essential feature of odor perception and description and is retrieved at each step of the neural processing of odors (Mantel et al., 2019). Together with hedonic valence, trigeminal dimensions seem to be another important parameter, as shown by Licon and colleagues (Licon et al., 2018). With a linguistic approach, Boisson (1997) studied odor naming in 60 language and found that hedonics are a very important part of odor naming in all of them, followed by intensity, temporal dynamics and moral virtue. David (2002) also retrieved this predominance of hedonics, as well as source-based descriptions. Her research showed that odor categorization is based on their source, intensity, memories, hedonics and use. Also, compared to vision and audition, she highlights the high prevalence of subjective markers in individuals' discourses.

These accounts on olfactory perception and language emphasize the importance of some perceptual and semantic parameters, but up today, there is no descriptive model that has been proposed that encompasses all the facets of perception. Here, we set out to construct such extensive descriptive model of olfactory perception from individuals' descriptions. To measure perception in less biased way and to explore facets that have been let aside, we developed a semi-guided interview adapted to olfaction. This approach is based on the *Explicitation Interview* (Petitmengin, 2006), that was designed to help individuals

verbalizing their subjective account of a phenomenon. The key word of this methodology is “how”: how does one experience a haptic sensation (Obrist et al., 2013), a pathology (Micoulaud-Franchi et al., 2012), a meditative state (Petitmengin et al., 2019) etc. In the context of this study, the interview was aimed at understanding how individuals experience odors, rather than asking them to identify their source. From the participants’ discourse, we extracted invariants in structures of perception, but also individual specificities. To enable cross-languages comparison, we conducted the study in two Western languages, namely German and French. We assume here that by alleviating the biases related to experimentation, we will observe forms of perception that are complementary of classical quantitative ratings and semantic descriptors. An important aspect of our approach is that the collected data is heterogenous and complex within and between individuals. To better account for such diversity, we applied data mining algorithm to extract relevant information and improve the description of the model. We hypothesize here that among the variability observed in the discourse of the individuals, we will distinguish between groups that share a similar perception, and ultimately better understanding the variability and universality of olfactory processing.

2. Materials & Methods

Participants

This study followed the statement of the Declaration of Helsinki. It was approved by the Comite de Protection des Personnes in France (CPP 2017/55) and by the ethical committee at the Medical Faculty of the Technical University of Dresden in Germany. Formal verbal and written consent were obtained from all the participants.

In total, 74 participants were involved in the experiment. They were recruited from the Technical University of Dresden, Germany (37 participants), as well as in Lyon, France (37 participants). The inclusion criteria were the following: no olfactory-related disorders, no neurological or psychiatric disorders, fluency in German or in French, age 18 to 50. All participants took a psychophysical test (Sniffing Sticks for the Germans, Hummel et al, 1997; European Test of Olfactory Capabilities for the French, Thomas-Danguin et al, 2003; providing a score from 0 to 16), in order to assess their olfactory abilities and screen for participants with olfactory impairment.

The mean age was 24 +/- 6.1 years old (53 women). Among the participants, 12 were often exposed to odors in their professional activity, 22 were sometimes exposed and 40 were not exposed. 7 participants were smokers.

Stimuli

A total of 20 odorants were used in this study. They were selected from the results of a pilot experiment, in which individuals were invited to freely describe 55 odorants. Their descriptions were very diverse and referred to visual images, sounds, textures, tastes, memories, irritation, freshness, warmth, well-being, stimulation, disgust, sensuality and

relaxation. They also rated these same odorants on pleasantness and intensity, as well as to what degree the characterization of the odorants was difficult, and the certitude in their identification if they proposed one. An experiential space of the 55 odorants was constructed with the answers of the participants. From this space, we selected the 20 most representative odorants, both perceptually and chemically. The selected odorants are described in the **Table 1**. Each odorant was diluted in mineral oil to obtain a total volume of 1mL, and then conditioned in a 15mL opaque vial (height 57mm, diameter, 27mm). The concentrations correspond to the concentrations necessary to get 1Pa of pressure in the air.

Code	Odorant	Concentration	CID
VAL	Isovaleric acid	0,48%	10430
BUT	Butyric acid	0,11%	264
BUTB	Butyl butyrate	0,18%	7983
PROP	Propanoic acid	0,04%	1032
ESAL	Ethyl salicylate	5,49%	8365
OCTN	1-octen-3-one	0,27%	61346
ACE	Acetic acid	0,01%	176
CIS	Cis-3-hexenol	0,24%	5281167
GUA	Guaiacol	2,09%	460
EUG	Eugenol	13,45%	3314
GER	Geraniol	11,29%	637566
BENZ	Benzaldehyde	2%	240
CIN	Cineole	1%	2758
LIN	(+)-linalool	2,17%	6549
BCAR	Beta-caryophyllene	33,64%	5281515
CAR	L-carvone	1,93%	439570
IAA	Isoamyl acetate	0,03%	31276
SAN	Alpha-santalol	Pure	5368797
MYR	Myrcene	0,15%	31253
EOCT	Ethyl octanoate	1,71%	7799

Table 1. Characteristics of the 20 odorant stimuli. The columns correspond to: code used to refer to the odorant in the analysis, common name of the odorant, concentration, Compound ID number in PubChem.

Protocol

Participants were informed of the main objectives of the study both orally and in a document written in their native language (French or German). After time to think, they gave or not their consent to be included in the study. People who agreed to participate received the instructions of the protocol and had the opportunity to ask questions before beginning.

In a first step, participants filled in a questionnaire including demographics items (age, sex, handedness, field of work), as well as elements related to their sense of smell: if they were

exposed to odors at work (Yes Often / Yes Sometimes / No), if they smoked, if they had a smell-related illness, their hunger level at the time of the experiment, their level of expertise (NA / Beginner / Amateur / Expert, later recoded as a 0 to 3 scale) in several domains (Art, Cooking, Wine Tasting, Perfumes, Music), if they practiced meditation (Yes Often / Yes Sometimes / No) and the importance they gave to each of their five senses in everyday life (from 1 to 9). They also took the Five Facets Mindfulness Questionnaire (Baer et al, 2006), which aimed to assess the individuals' capacity to pay attention and describe their internal and external environment, especially at the sensory level. This questionnaire was a short version with 15 items (Baer et al, 2012), providing a global score ranging from 5 to 75, with five sub scores corresponding to the different facets of mindfulness (Observing, Describing, Acting with awareness, Non Judging and Non reacting).

Secondly, participants received specific written instructions about the interview, in their native language, as the following: "In this study, we are interested in what manifests to you when you smell odors. To do so, imagine the following situation: a crime happened, and a detective arrives to interrogate the witnesses. During the interviews, he asks them to describe what they saw or heard at the time of the crime. He does not want the witnesses to make assumptions on what happened, only to explain as precisely as possible their perception at this moment. Here you are in the situation of a witness. We want to hear your subjective perception, and not an interpretation on the source of the odors that will be presented to you. As a result, there is no right or wrong answer. For example, if you think that a vial smells like a rose, the fact that the odor typically comes from a rose or not is not important in this experiment. We are more interested in the fact that, for you, this odor is associated to a rose. We know that this descriptive exercise can be difficult, but we ask you, as far as possible, to pay attention to the odors that you will be presented with and what they evoke for you, and to let aside your external preoccupations."

Then, the participants were presented with 20 vials containing odorants. They also had a training vial to make sure that the instructions were understood. For each trial, the participants smelled a vial by themselves as many times they needed and described their subjective experience of the odorant within 2 minutes, according to a short semi-guided interview conducted by the experimenter. If the participants spontaneously talked about their experience, the interviewer would not interrupt them. When they had difficulties or stopped talking, the experimenter asked questions aimed to stimulate verbalization: "*What does this odor evokes for you?*", "*What do you think when you smell this odor?*", "*If you had to describe this odor to a friend who has never experienced it, what would you say?*". The entire interview was audio recorded for later transcription of the discourse.

After each vial, the subject rated the difficulty they had to verbally describe the odor (Not difficult at all / A little bit difficult / Moderately difficult / Very difficult / Extremely difficult, later recoded as a 1 to 5 scale). This whole interview was audio recorded to allow for the transcription of the verbalizations of the participants. Afterwards, participants were to rate each odorant on a quantitative scale (from 1, *not at all*, to 9, *very*) on the following

dimensions: intensity, familiarity, edibility, hedonic value (the scale here was: -4, *very unpleasant*, to 4, *very pleasant*) and irritation.

Data analysis

We report how we determined our sample size, all data exclusions (if any), all manipulations, and all measures in the study, and we follow JARS (Kazak, 2018). All data, analysis code, and research materials are available at [stable link to repository]. Data were analyzed using R, version 4.0.0 (R Core Team, 2020) and the package ggplot, version 3.2.1 (Wickham, 2016). This study's design and its analysis were not pre-registered. In the subsequent analysis, we separated measurements of hedonic valence and emotional saliency. To this end, verbalizations referring to pleasantness were separated into positive and negative references. In addition, pleasantness ratings were also coded in terms of emotional saliency, corresponding to the absolute distance to 0.

Secondly, when we investigated the effect of sex, exposure to odors and language on the verbalizations of the participants, we ensured that confounding variables were controlled in the samples of comparison. Indeed, for the effect of sex, we first filtered the data to extract individuals who were not exposed to odors (13 men and 27 women). In order to have equivalent groups, we then randomly selected 13 women in this sample. A Kruskal-Wallis test was conducted to compare verbalizations in the 6 main categories according to sex. A total of 999 replications was performed and the mean results were used to interpret the test. The same pipeline was used for the effect of language: we first filtered the data to extract individuals who were not exposed to odors (28 French and 12 Germans) and then randomly selected 12 French in this sample for the subsequent test. Finally, for the effect of exposure to odors, we first filtered the data to extract females (27 never exposed and 26 sometimes or often exposed), and then randomly selected 26 in the non-exposed group for the subsequent test.

3. Results

The subjective experience of smells is structurally organized

To investigate the structure of the olfactory perception, a total of 74 participants from two linguistic backgrounds (37 French and 37 Germans) were asked to describe their subjective experience of 20 representative odorants, through a semi-guided interview. An important facet of this paradigm was that the instruction was not to identify the odor per se nor to use semantic descriptors, but rather to express in their own words what it was like for them to smell the odors.

After transcription and translation, the participants' verbalizations were encoded according to the references they made, which could be a concept, an object, a memory or a quality (see **Methods**). 206 distinct references were encoded in total. Participants used in average 9.16 +/- 2.47 references to describe an odor.

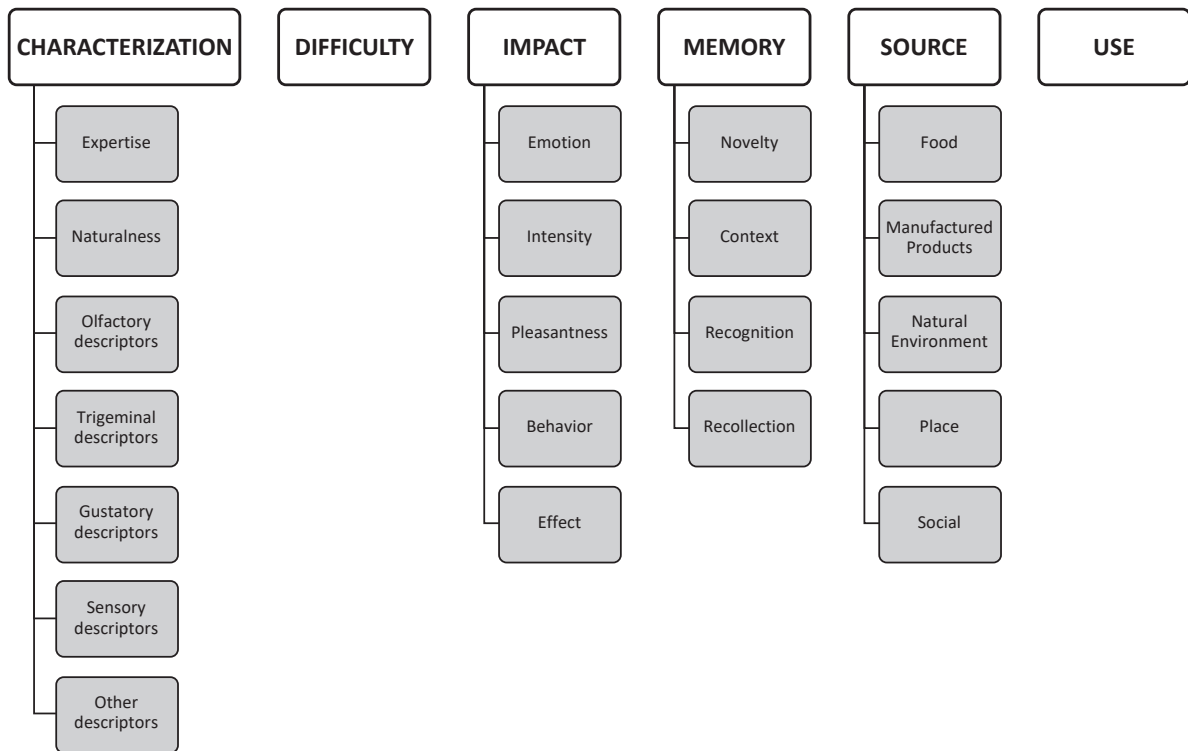


Figure 1. Descriptive model of the subjective experience of odors.

From the encoded references, a descriptive model of the subjective experience of odors was then constructed (see **Figure 1**). This model was composed of 6 main categories and 24 subcategories of references, described in **Table 1**.

It can be noted that such categories were not used equally by the participants: *SOURCE* (31.3% of references), *IMPACT* (27.8%) and *CHARACTERIZATION* (25.4%) constituted the majority of the verbalizations, whereas individuals used less references to *MEMORY* (9%), *DIFFICULTY* (4.1%) and *USE* (2.5%). We further looked for the substructure of these categories. This showed that the most prevalent were *Food* (within *SOURCE*, representing 14.2 % of global references) and *Pleasantness* (within *IMPACT*, 12.8% of global references), as shown in **Figure 2a**. The least used subcategories were *Expertise* (0.5% of references), *Other Source* (0.6% of references) and *Novelty* (1,1% of references). **Figure 2b** shows the proportion of references in each main category, with the addition of the two most used subcategories.

Category	Subcategory	Content	Examples
IMPACT	Emotion	Emotions elicited by the odor	"joy", "wellness", "nostalgia", "disgust"
	Intensity	Strength of the odor	"intense", "low", "light", "discrete"
	Effect	Effect of the odor on the individual	"stimulating", "relaxing", "surprising"
	Pleasantness	Positive or negative value associated to the odor	"pleasant", "negative", "horrible"
	Behavior	Approach or avoidance behavior elicited by the odor	"I want to leave the room", "I could smell this all day"
SOURCE	Food	Everything edible for humans	"dairy", "fruits", "spices", "vegetables", "meat", "sweets"
	Manufactured products	Products fabricated by humans and industrial materials	"medication", "book", "soap", "car", "metal"
	Natural environment	Elements of nature	"flower", "soil", "wood", "grass"
	Places	Specific place where the odor can be encountered	"hospital", "garden", "flower store"
	People	Specific persons and human body odors	"my mother", "vomit", "urine", "transpiration"
	Other sources	Other references	"smoke", "fire", "mold"
CHARACTERIZATION	Expertise	Expert terms to describe the odor, its chemical name or characteristics	"butyric acid", "ester"
	Olfactory descriptors	Olfactory terms, abstract or source-based	"moldy", "rotten", "floral", "vegetal", "lemony", "fishy"
	Gustatory descriptors	Terms related to taste	"sour", "sweet", "salty", "pungent"
	Trigeminal descriptors	Terms related to the trigeminal features of the odor	"cold", "irritating", "painful", "spicy"
	Sensory descriptors	Other sensory terms, referring to vision, audition, texture...	"I can see the object", "green", "pink", "wet", "dry", "soft"
	Naturalness	Natural or artificial origin of the odor	"natural", "chemical", "synthetic", "artificial"
	Other descriptors	Other descriptors	"complex", "volatile", "rare"
MEMORY	Novelty	Odor not previously encountered	"I don't know this odor", "new"
	Recognition	Familiar odor not associated with a context	"I know this odor"
	Context	Familiar odor associated with a spatial or temporal context	"This is the odor when I am cooking", "an odor of Christmas"
	Recollection	Odor associated to a specific event or memory	"I smelled this odor during my trip to Italy"
DIFFICULTY		References to the difficulty for the subject to verbalize, describe or associate something with his/her experience	"I have no words for this", "I don't know how to describe", "I can't find the name of the odor"
USE		References to the use for the odor	"medical", "cooking", "edible", "it signals danger"

Table 2. Description and examples for each identified category of references and its subcategories.

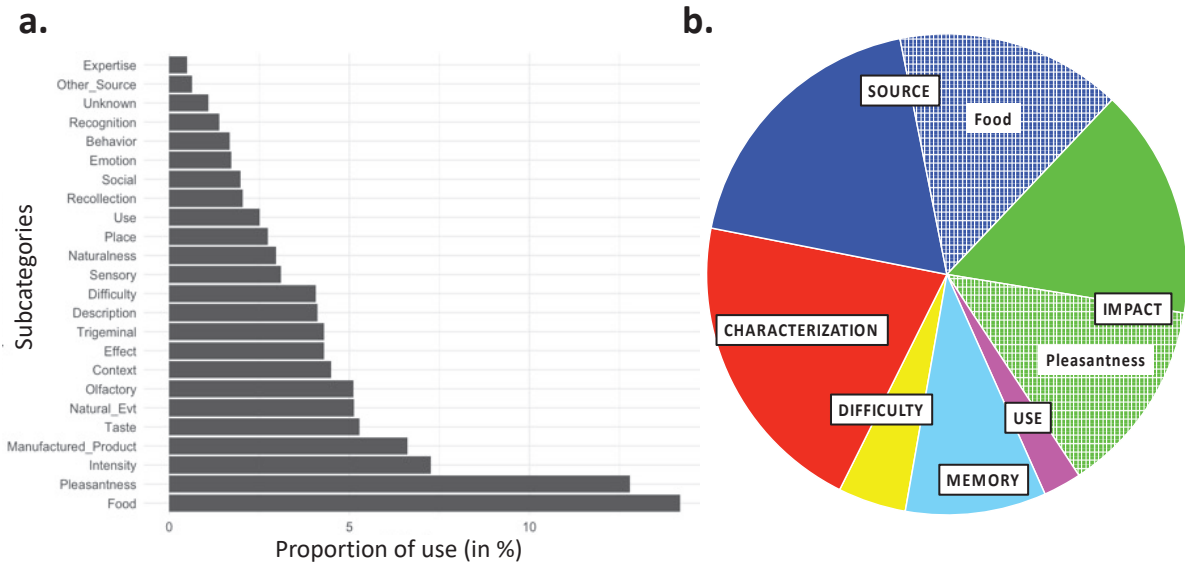


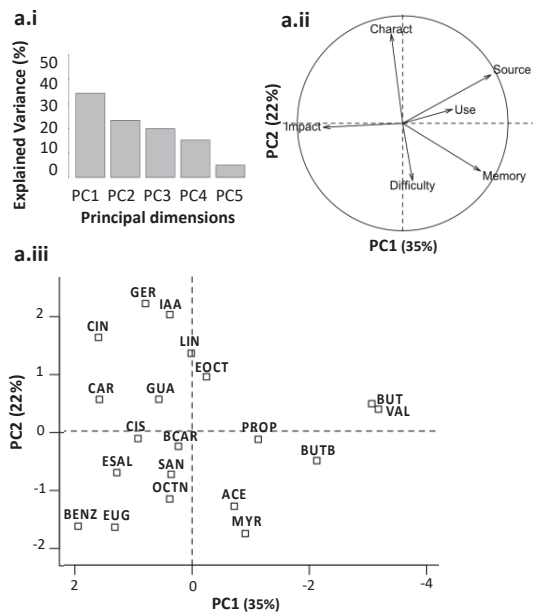
Figure 2. Use of categories and subcategories of references. *a.* Proportion of use for each subcategory. *b.* Proportion of use for each main category. The food and pleasantness subcategories are also indicated into the respective categories of Source and Impact.

To understand how the subjective space was structured, we performed a Principal Component Analysis on the verbalization categories across the 20 odorants (see **Figure 3a**). Results showed that the first two principal components (PC) explained 77% of the total variance. The first PC was correlated with references to *SOURCE* ($r = .84, p = .0001$), *IMPACT* ($r = -.75, p = .003$) and *MEMORY* ($r = .74, p = .005$). The second PC was correlated with references to *CHARACTERIZATION* ($r = .83, p < .001$) and *DIFFICULTY* ($r = -.86, p < .001$).

We further asked the same question for the subcategories and showed that the first three principal components explained 58% of the data (see **Figure 3b**). The first PC was correlated with references to *Behavior* ($r = .80, p = .003$), *Pleasantness* ($r = .75, p = .02$) and *Social* ($r = .90, p < .0001$). The second PC was correlated with references to *Food* ($r = .75, p = .02$) and *Novelty* ($r = -.86, p < .001$). The third one was correlated with references to *Taste* ($r = -.77, p < .01$) and *Trigeminal* ($r = .73, p = .03$).

In sum, verbal descriptions of olfactory experience are structurally organized, and are mostly constituted of references to the source of the odor, its impact on the individuals and its qualitative description. More specifically, food-related terms and hedonic value are the main elements of people's verbalizations.

a. Categories



b. Subcategories

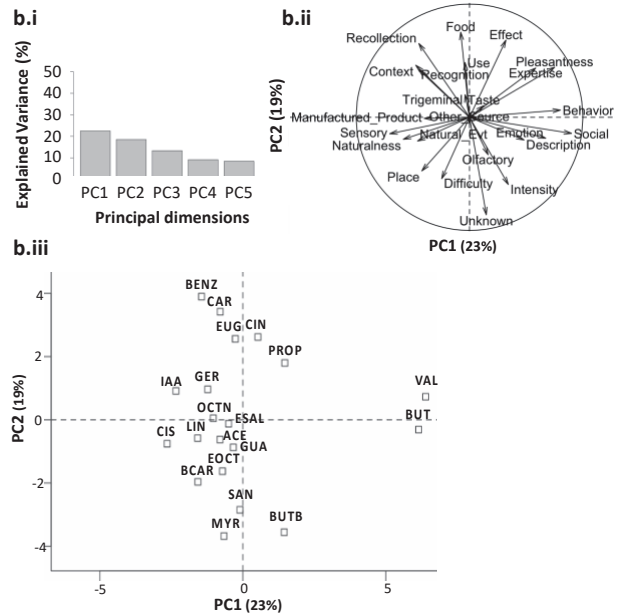


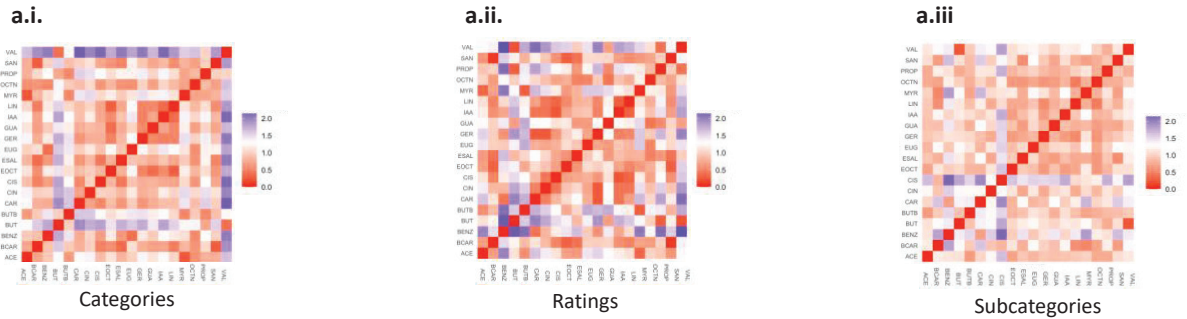
Figure 3. Results of the Principal Components Analysis for the verbalization spaces. a.i. Histogram of the percentage of explained variance for each principal component extracted from the PCA analysis on categories. a.ii. Correlations between each of 6 categories with $PC1_{Categories}$ on the x-axis and $PC2_{Categories}$ on the y-axis. a.iii. Representation of the categorial space showing the 20 odorants in a 2D space with $PC1_{Categories}$ on the x-axis and $PC2_{Categories}$ on the y-axis, as well as their explained variance (%). b.i. Histogram of the percentage of explained variance for each principal component extracted from the PCA analysis on subcategories. b.ii. Correlations between each of 24 subcategories with $PC1_{Subcategories}$ on the x-axis and $PC2_{Categories}$ on the y-axis. b.iii. Representation of the categorial space showing the 20 odorants in a 2D space with $PC1_{Categories}$ on the x-axis and $PC2_{Categories}$ on the y-axis, as well as their explained variance (%).

The subjective experience and standard quantitative ratings of smells share some features but have their own specificities

In order to ask how the subjective experience of smells relates to standard quantitative measurements of odors, we further asked our participants to rate the same stimuli along five perceptual dimensions (intensity, irritation, pleasantness, familiarity and edibility) using visual scales (see Methods). First, the number of references in each verbalization subcategory was correlated with the perceptual ratings of each odorant. Intensity ratings were significantly linked with references to the subjective subcategories of *Pleasantness* ($r = .71, p < .001$), *Effect* ($r = .76, p < .001$) and *Sensory Descriptors* ($r = -.72, p < .001$). They also had a quadratic relationship with references to *Intensity* ($F(2,17) = 12.3, p < .001$). Familiarity ratings were correlated with references to *Intensity* ($r = -.80, p < .001$), *Novelty* ($r = -.85, p < .001$), and *Recollection* ($r = .79, p < .001$). Hedonic ratings were related to references to *Intensity* ($r = -.74, p < .001$) and had a quadratic relationship with references

to *Pleasantness* ($F(2,17) = 9.7, p < .01$). Finally, difficulty ratings were associated with references to *Novelty* ($r = .88, p < .001$).

a. Distance matrices of odorants in the different spaces



b. Correlations between spaces

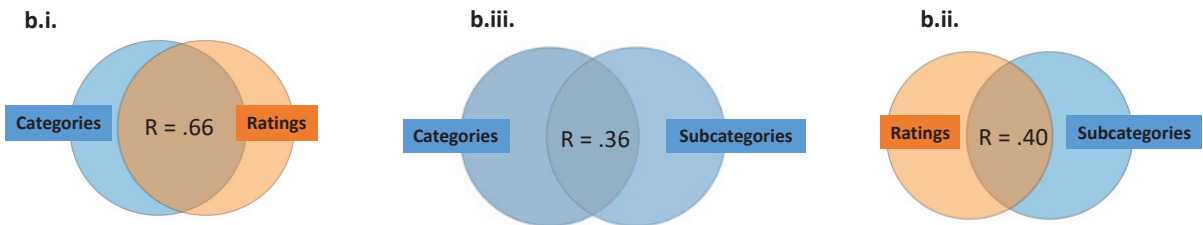


Figure 4. Correlation between the ratings space and the verbalizations spaces. *a. Matrices of normalized Euclidian distances between the odorants for the categorical (a.i), ratings (a.ii) and categorial space (a.iii). The scale indicates the Euclidian distance values divided by the mean for each matrix. b. Correlations between the spaces. The Venn diagrams represent the overlap of the spaces as shown by the Mantel correlation between the two distance matrices, namely categorical vs ratings (b.i), categorical vs subcategorical (b.ii) and ratings vs subcategorical (b.iii).*

In a second step, we compared the perceptual ratings and the verbal descriptions to determine if these two aspects of perception were associated. To this end, we built Euclidean distance matrices between the odorants and computed a Mantel test with 999 permutations. It showed that the ratings and sub categorical spaces of odorants were significantly correlated ($r = .41, p = .001$), as well as the ratings and categorical spaces of odorants ($r = .66, p = .001$), as displayed in **Figure 4**.

Taken together, these results show that, while the verbalizations and rating spaces of odorants have similarities, verbalizations bring additional information compared to standard ratings, especially in terms of qualitative description, difficulty to describe and reported usage of the odors.

The subjective experience is structured differently from one odorant to another

In a third series of analyses, we then asked how the structure of subjective experience shifts according to the type of stimuli. To this end, we applied the same pipeline of analysis for each specific odorant. We first show that odors significantly differed in their usage of the references to *SOURCE* ($X^2_{\text{Friedman}}(19) = 124.3, p < .001$): individuals referred more to this category for CIS, BENZ, CAR, CIN, ESAL, GER compared to other odors. They also significantly differed in the references to *IMPACT* ($X^2_{\text{Friedman}}(19) = 104.8, p < .001$): individuals referred more to it for BUT and VAL. Regarding *CHARACTERIZATION*, there was also a significant difference ($X^2_{\text{Friedman}}(19) = 52.6, p < .001$), with more references to it for IAA. For *MEMORY*, it was also the case ($X^2_{\text{Friedman}}(19) = 122.7, p < .001$), with more references for BENZ, EUG and CIS. For *USE* ($X^2_{\text{Friedman}}(19) = 76.3, p < .001$), there were more references for CIN compared to BUTB only. Finally, for *DIFFICULTY* ($X^2_{\text{Friedman}}(19) = 57.6, p < .001$), there were no significant differences after post-hoc comparisons. **Figure 5** shows the proportion of references in categories for each odorant.

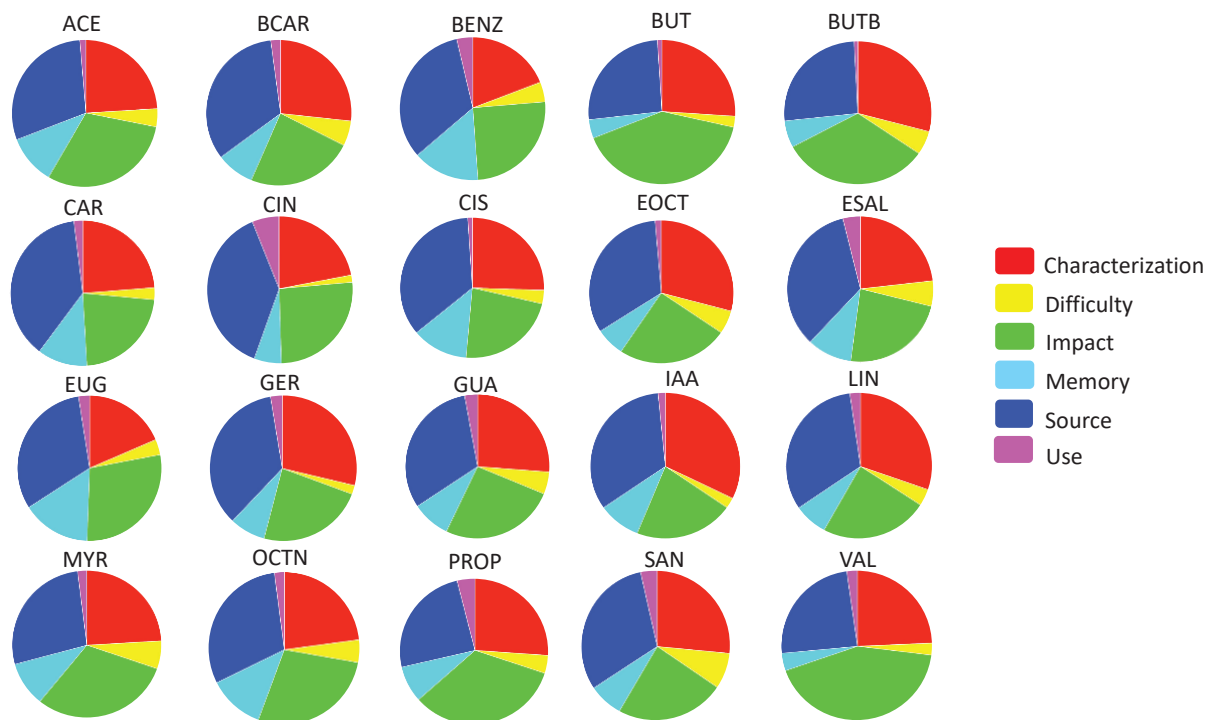


Figure 5. Proportion of use of each category of references across odorants

In sum, the descriptive model shows stability across different types of stimuli even though the relative proportion of categories can be modulated depending on their nature.

Data mining analysis provides fine-grained descriptions of the subjective experience across individuals

Olfactory perception varies greatly from one individual to another, and this depends on a series of factors, such as sex (Brand & Millot, 2001; Sorokowski et al., 2019), age (Doty, 2018; Doty & Kamath, 2014), level of expertise (Royet et al., 2013), culture (Ferdenzi et al., 2013; Majid, 2015) or hunger state (Albrecht et al., 2009). Here, we explore how three prominent factors of variation, namely sex, language and exposure to odors influenced the structure of the subjective experience of smells.

In a first standard analysis, we showed for sex, that there was a significant difference for *Characterization* ($X^2_{KW}(1) = 4.34, p = .048$): women used more references to this category compared to men ($M_{women} = 2.20 \pm 0.88, M_{men} = 1.36 \pm 0.98$). For language, we did not show any significant difference between French and German. With regards to exposure, results showed that there was a significant difference for *Source* ($X^2_{KW}(1) = 7.13, p = .008$): individuals exposed to odors used more references to this category compared to the ones who did not ($M_{exposed} = 3.23 \pm 1.01, M_{not_exposed} = 2.54 \pm 0.92$). It was also the case for *Characterization* ($X^2_{KW}(1) = 5.04, p = .03, M_{exposed} = 2.86 \pm 1.11, M_{not_exposed} = 2.20 \pm 1.15$).

In a second step of analysis, we used a data mining technique named *subgroup discovery* through the *MTCSExtent* algorithm (Mathonat et al., 2021). This method allows to identify some patterns that are over-represented in a subgroup (e.g., men) and under-represented in another (e.g., women). We focused here on the *informedness* parameter, which corresponds to the difference between the true positive rate and the false positive rate of a given pattern in the subgroup. The higher the value, the more specific the pattern to such subgroup. The best patterns found for each subgroup of interest are represented in **Figure 6**. Regarding sex, men used the descriptor “Food” 4% more compared to women, whereas women mentioned “Fresh” 7% more compared to men. In general, we see that the descriptors that are more used in men mainly refer to *Source* (“Food”, “Plant”) and *Impact* (“Strong”, “Discomfort”, “Positive”, “Aggressive”, “Appetizing” and “Negative”) whereas for women, the most specific descriptors refer to *Characterization* (“Fresh”, “Sweet”, “Acrid”, “Soft”, “Fresh”, “Chemical”). Regarding language, French participants use “Strong” 11% more than Germans, but German individuals used “Acrid” 10% more compared to French. In general, the French tended to use more descriptors belonging to *Impact* (“Strong”, “Soft”, “Aggressive”, “Surprise”) and Germans to *Characterization* (“Acrid”, “Artificial”, “Acidic”, “Sweet”, “Fresh”). Finally, regarding exposure to odors, participants exposed to odors in their life used “Sweet” 11% more than those who were not exposed, whereas the latter mentioned “Difficulty to identify” 10% more than exposed individuals. In general, people exposed to odors used more descriptors belonging to the *Characterization* category (“Sweet”, “Acidic”, “Acrid”, Fruity”, “Floral”, “Artificial”) whereas the ones who were not used more descriptors from the *Difficulty* (“Difficulty to identify”, “No evocation”).

In summary, the pattern of verbalizations differed according to several individual factors of variation and shows that some elements bear more significance for some groups of individuals.

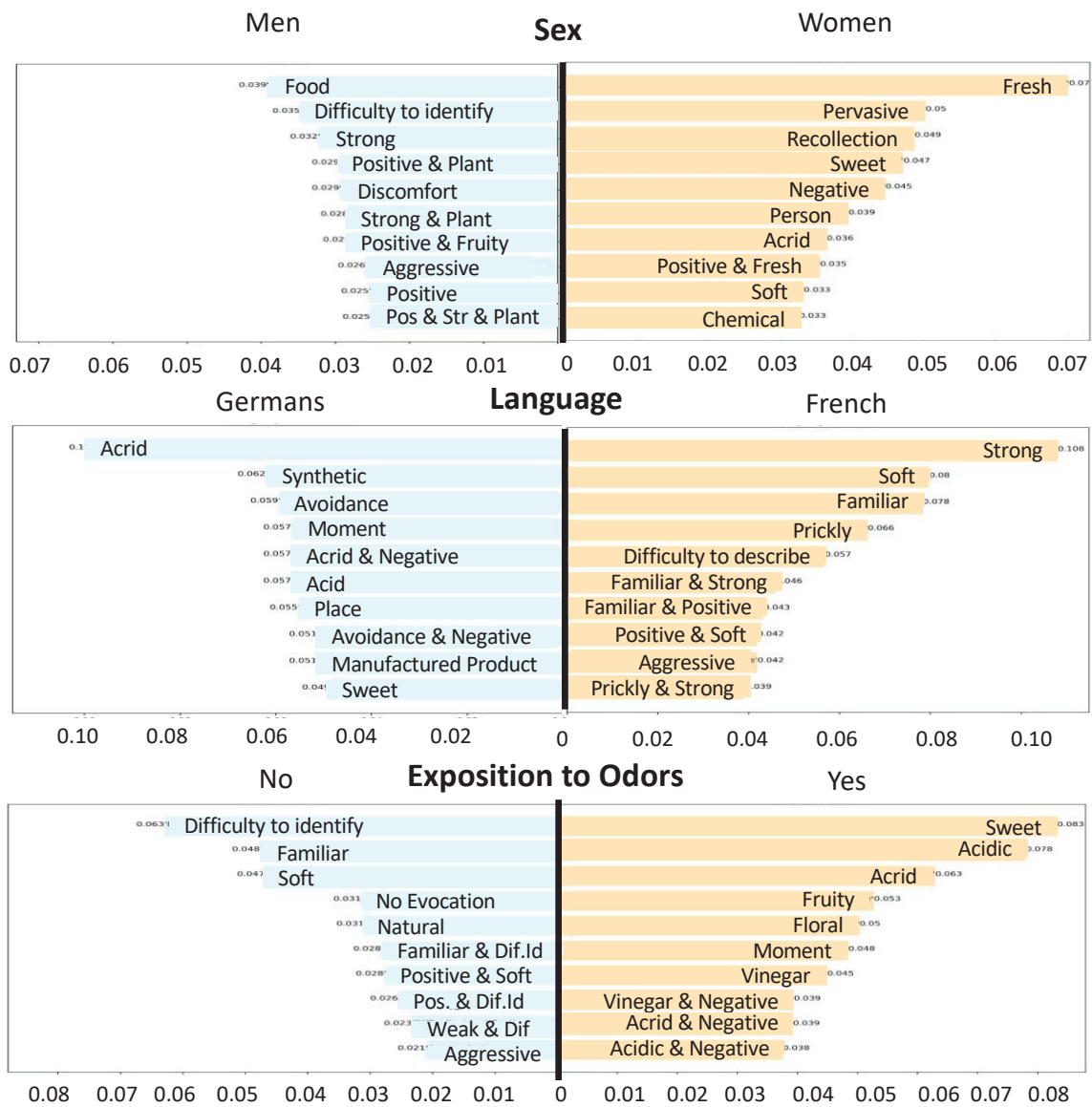


Figure 6. Subgroup discovery analysis on the groups of individuals. Best patterns of descriptors and associated informedness for Sex (a), Language (b) and Exposition to odors (c). Informedness of the patterns are represented on the x-axis, corresponding to the ratio between the true positive rate and the false positive rate of occurrence of specific descriptors. The ten best patterns of descriptors are represented on the y-axis for each group of individuals (top to bottom: best to worst). If two descriptors are indicated on the same line, this means that the pattern corresponded to the concomitant use of these descriptors.

Clustering analysis reveals different sub-groups of individuals sharing similar subjective experience

Based on the above findings showing variations depending on odorants and individuals, we then explore the fine-grained specificities of the identified subgroups of individuals responding to each odorant. To this end, we performed a hierarchical clustering with Ward distances for each of the twenty odorants (with R packages “stats” and “factoextra”). The optimal number of clusters was chosen according to the highest mean value of the silhouette width, with a maximum of 10 clusters. The results of the clustering are represented in **Figure 7**. Overall, for each odorant, two clusters of individuals were identified with a quite good quality of clusterizing, as the mean silhouette width was between 0.53 and 0.59.

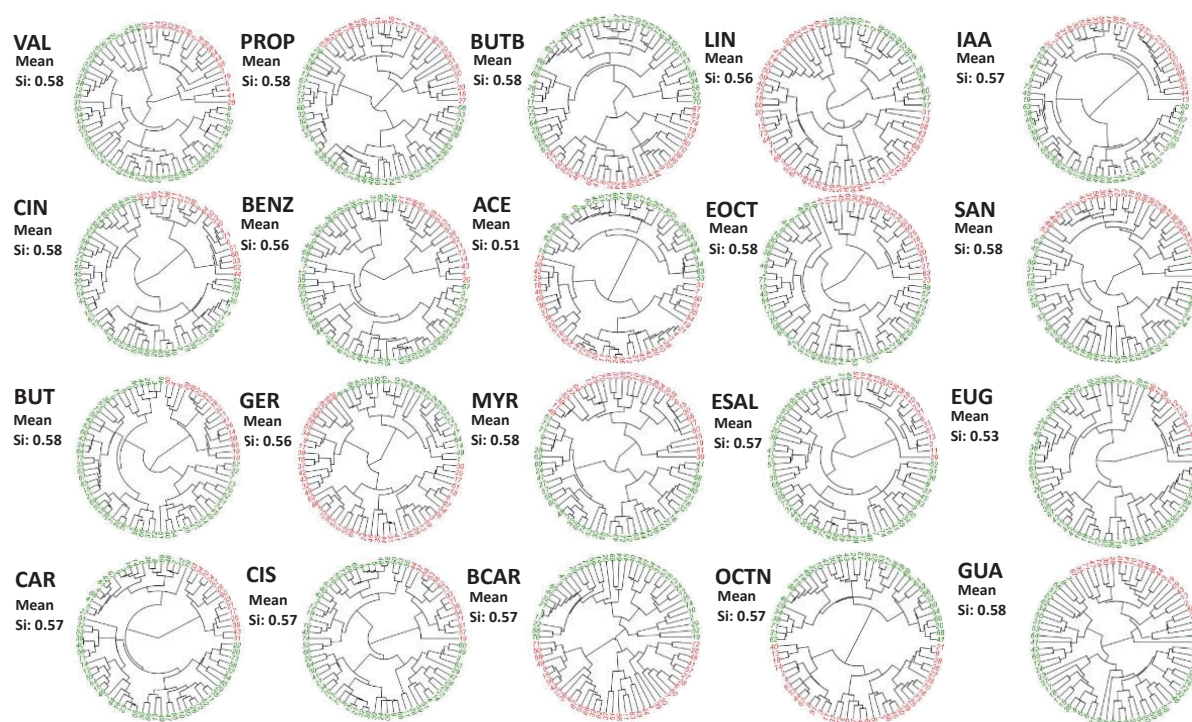


Figure 7. Hierarchical clustering of individuals for each odorant. The mean silhouette width is indicated for each odorant.

In a next step of analysis, we used another technique of data mining, in order to characterize the identified clusters of individuals for each odorant. The aim was to characterize each cluster with the most frequent and exceptional types of references that were used by the participants. To this end, we used the *a priori closed algorithm* (Pasquier et al., 1999) to identify frequently used patterns within the verbalizations. Among these, we selected the exceptional itemsets by keeping the patterns that were at least twice as frequent in a given cluster compared to the other. Then we retained the minimum of patterns characterizing the maximum of subjects within this cluster.

Odor	Cluster 1				Cluster 2			
	Cluster Size (Coverage)	Descriptors	Cumulated Coverage	Growth Rate	Cluster Size (Coverage)	Descriptors	Cumulated Coverage	Growth Rate
ACE	34 (100%)	Vinegar Negative, Pleasant Food	88% 97% 100%	7.1 7.0 3.8	40 (90%)	Weak Difficulty to Identify Avoidance	65% 80% 90%	3.7 12.8 2.3
BCAR	38 (82%)	Place Weak Manufactured Product Inside, stuffy place Negative, Unpleasant	34% 56% 74% 79% 82%	4.1 3.5 4.7 Inf. Inf.	36 (100%)	Positive, Pleasant Plant Wood, Forest Natural Fresh	56% 89% 94% 97% 100%	2.6 4.7 2.1 13.7 10.6
BENZ	57 (98%)	Almond Pastry Sweet	86% 96% 100%	7.3 2.7 2.6	17 (100%)	Familiar Difficulty to Identify Recollection Glue	71% 88% 94% 100%	5.0 3.3 2.9 6.7
BUT	56 (57%)	Avoidance Comparison Disgust	32% 48% 57%	Inf. 4.2 4.8	18 (100%)	Cheese Strong	78% 100%	6.2 2.1
BUTB	45 (73%)	Negative, Unpleasant No Evocation	62% 73%	6.0 5.8	29 (90%)	Sweet Positive, Pleasant Soft Fruit Fruity	45% 72% 83% 86% 90%	Inf. 15.5 3.1 8.5 4.7
CAR	60 (100%)	Mint Positive, Pleasant Gum Strong	75% 96% 98% 100%	10.5 3.2 4.7 3.3	14 (86%)	Familiar Candy Difficulty to Identify	64% 79% 86%	7.7 8.6 15.0
CIN	52 (87%)	Positive, Pleasant Mint	75% 87%	2.4 9.7	22 (100%)	Medical Product Eucalyptus	83% 100%	4.5 3.9
CIS	59 (29%)	Natural	29%	2.2	15 (100%)	Grass & Moment Fresh	80% 100%	15.7 6.2
EOCT	51 (70%)	Weak Negative, Unpleasant Difficulty to Describe	43% 61% 70%	2.5 6.3 2.9	23 (100%)	Fruit Sweet Soft	78% 96% 100%	5.0 5.4 8.1
ESAL	55 (75%)	Positive, Pleasant Place Familiar	47% 65% 75%	Inf. 2.6 4.5	19 (100%)	Negative, Unpleasant Manufactured Product	95% 100%	7.4 5.0
EUG	60 (75%)	Strong Negative, Unpleasant	52% 75%	2.4 Inf.	14 (100%)	Clove	100%	2.1
GER	26 (88%)	Citrus Fresh	73% 88%	3.2 2.8	48 (94%)	Positive, Pleasant Sweet Plant	90% 92% 94%	7.8 2.0 Inf.
GUA	48 (69%)	Positive, Pleasant Recollection Familiar	40% 56% 69%	10.3 Inf. Inf.	26 (96%)	Negative, Unpleasant Acrid Prickly	88% 92% 96%	4.7 9.2 4.9
IAA	45 (100%)	Positive, Pleasant Fruit Sweet Candy	76% 93% 98% 100%	4.4 Inf. 2.4 16.8	29 (83%)	Place Artificial, Synthetic Moment Negative, Unpleasant	41% 62% 76% 83%	9.3 2.8 2.2 3.1
LIN	21 (100%)	Flower Soft Floral	76% 95% 100%	13.5 4.6 3.8	53 (57%)	Citrus Fresh Moment	28% 45% 57%	Inf. Inf. 2.2

MYR	43 (86%)	Weak Difficulty to Identify Wood, Forest No Evocation	49% 70% 81% 86%	2.2 2.5 Inf. 3.4	31 (97%)	Manufactured Product Negative, Unpleasant Place	61% 87% 97%	Inf. 3.6 7.6
OCTN	38 (87%)	Mushroom Positive, Pleasant Wood, Forest Recollection	61% 79% 84% 87%	21.8 4.3 6.6 6.6	36 (89%)	Negative, Unpleasant Familiar Difficulty to Describe Manufactured Product	58% 78% 86% 89%	3.2 12.7 8.4 5.3
PROP	45 (89%)	Vinegar Prickly Acidic	76% 84% 89%	22.0 2.4 3.4	29 (83%)	Avoidance Difficulty to Identify Difficulty to Describe Negative, Unpleasant Chemical	38% 62% 72% 79% 83%	3.4 5.2 3.1 4.7 12.4
SAN	43 (91%)	Positive, Pleasant Perfume Person Sweet Soft	72% 81% 86% 88% 91%	Inf. Inf. Inf. Inf. 2.2	31 (42%)	Negative, Unpleasant	42%	2.6
VAL	53 (53%)	Disgust Pervasive Place	26% 43% 53%	2.8 2.6 Inf.	21 (100%)	Cheese Body Odor Positive, Pleasant	86% 95% 100%	7.6 2.2 8.8

Table 2. Cluster characterization of the 20 odorants. The cluster coverage corresponds to the percentage of individuals who used the combination of all the descriptors identified for the cluster. The cumulative coverage corresponds to the cumulative percentage of individuals who used the particular descriptors within the cluster. The growth rate corresponds to the ratio of the prevalence of the descriptor within the cluster upon the prevalence outside the cluster. The higher this indicator, the more the pattern can be considered as specific to the cluster.

With this method, we were able to identify patterns of responses that characterized the majority of individuals within the clusters. For example, for the odorant ACE, 100% of the individuals of the first cluster mentioned the descriptors “vinegar”, “unpleasant” or “food”, and 90% of those in the second cluster used the descriptors “weak”, “difficulty to identify” or “avoidance”. For the odorant BENZ, 98% of the individuals in the first cluster mentioned “almond”, “pastry” or “sweet” and 100% of those in the second cluster used the descriptors “familiar”, “difficulty to identify”, “recollection” or “glue”. Overall, 13 odorants out of 20 had a coverage superior to 80%, which means that the description covers most individuals. However, not all clusters were as precisely characterized. For CIS for example, although 100% of individuals used “grass”, “moment” or “fresh” to describe the odorant in the first cluster, the second cluster was far more diverse, with a coverage of only 26% of individuals who mentioned “natural”. It was also the case for some groups of BUT (57% coverage), EUG (75%), GUA (69%), LIN (57%), SAN (42%) and VAL (53%) odors. The details of the clusters’ coverage and descriptors usage are represented in **Table 2**.

In summary, a data mining analysis allowed to identify and characterize subgroups of individuals with similar responses for each odorant, which differed on various categories depending on the stimuli. Interestingly, the subgroups had different compositions from an odorant to another, suggesting that variability in responses is driven by the type of stimulus in addition to inter individual diversity.

4. Discussion

The aim of the present study was to better understand the cognitive underpinnings of the subjective experience of odors according to a series of prominent factors of variation, including gender, language, and exposure to odors. We first showed that the subjective experience of smells is structurally organized. In addition, although they share some features with standard quantitative ratings, verbalizations have their own specificities and constitute a complementary approach for the measurement of perception. We also found that subjective experience of odors is structured differently depending on odorants and individuals. To further explore these variations, data-mining algorithms provided fine-grained descriptions of this experience among individuals and a clustering analysis allowed to characterize different subgroups of individuals sharing similar experience across the odorants.

From these results, the first element of interest was that contrary to common belief, people were able to talk about odors, with an average of 9 references per odorant. This can be due to the task that was asked to the participants, in which they had to describe what it was like for them to smell odors and not to identify or characterize the odor per se as it is often the case in experimental research on olfaction. They were thus able to mention various elements besides the source or olfactory descriptors: the impact the odors had on them, memories associated to the odors, the difficulty they had to identify or describe the odors, and the usage they would make of the odors. When characterizing the odor per se, they not only used common olfactory descriptors (i.e., floral, fruity...) but also borrowed terms from other sensory modalities, especially gustatory or trigeminal terms. They talked also about the naturalness of the odors. These different elements have been studied in the olfactory community, but often in isolation. Indeed, some studies were interested in the characterization of olfactory qualities: for example, Dravnieks's atlas (Dravnieks, 1985) comprises a set of 146 terms that can be applied to odors. Different kinds of measurement for odor qualities have been developed, such as sorting, classification or ratings of similarity (Wise et al., 2000). Furthermore, although experimental research has shown that the source of odors is often mentioned when people are asked to describe them, odor naming is known to be especially hard (Cain, 1979), at least for Westerners (Majid & Burenhult, 2014). Other studies focused on the propension of odors to evoke memories, especially autobiographical (Larsson & Willander, 2009), and the difference between negative and positive odors (Larsson et al., 2009). Perceptual aspects of odors are also investigated, especially pleasantness (Distel et al., 1999; Rouby et al., 2009). The present experiment allowed us to build a model including these aspects present in the scientific literature, but with new information of the relative importance of the different dimensions in spontaneous verbalizations. We thus showed the overwhelming importance of hedonic and food-related terms, consistent with the main functions of smell in humans, namely identification of food or potential hazards.

We also investigated how common individual factors of variation affect the descriptive model. Gender had an influence on the type of descriptions that were used, with more qualitative characterization of the odors by women. There is a consistent but small effect of sex on odor perception (Sorokowski et al., 2019), as women tend to have better performances in odor detection, identification and discrimination. Here, we show that there is also a qualitative difference in how women describe their experience of odors compared to men.

Also, exposure to odors was an important factor in verbalizations. People who were frequently or sometimes exposed to smells at work used more references to source and to qualitative aspects of the odors. We know that experts in olfactory-related fields have better verbal abilities to describe odors (Royet et al., 2013; Sezille et al., 2014), but here we show that there is a difference even for less related professions such as healthcare.

Interestingly, language had no influence on the types of references individuals used. Here, we compared French and German, two countries with similar cultural and culinary practices, and differences may be too little to be identified. It has otherwise been shown that hunter-gatherers (Majid & Kruspe, 2018) especially have extensive smell vocabularies and are able to talk easily about odors, so it is reasonable to expect a different organization of smell experience for other populations. Future studies will be necessary to better understand how the model behaves for diverse cultures and languages.

Finally, we used a data-mining based approach to better characterize the individuals in terms of their verbalizations, and to be able to identify subgroups sharing similar perceptions. We found that for all the tested odorants, two main groups of people could be characterized. These groups were each focused on different aspects of the odors: for example, for odorant BENZ, some people were describing the source of the odor, whereas the others mentioned more its familiarity and memories associated to it. This approach is very interesting to understand the organization of olfactory perception, and especially to make sense of the observed variability in peoples' verbalizations of odors. It shows us that although individuals seem to share a general system of perception, some pay more attention to specific aspects compared to others. This difference in focus is often invisible in classical quantitative ratings, as the experimenters guide the participants toward each element of interest, which erases the spontaneity of attention.

In conclusion, the present experiment is to our knowledge the first attempt to construct a descriptive model of odor perceptual experience, with less biased method of acquisition based on guided verbal description of subjective experience of smells. It gives us valuable information about the general structure of odor perception, and data mining analysis shows that the individual variability lies in the attention given to particular facets of perception. This model constitutes a basis for future experiments on subjective experience of odors and its variability according to specific factors of interest. More generally, it allows for a better understanding of the complex relationship between perception and language, in the context of olfaction.

Acknowledgements

The present study was funded by a grant from CNRS - LIA/IRP Human Chemosensation to MB

Authors' contributions

Conceptualization: MM, MB, JMR; Formal Analysis: AF, MMa, MMo, MP; Funding acquisition: MB, Investigation: MM, AO, JC; Methodology: MM, MB, AF, TH; Resources and Software: TH; Supervision: TH, MB, JMR, MP; Writing (Original Draft and Preparation): MM, MB; Writing (Review and Editing): All authors.

Competing interests

None

References

- Albrecht, J., Schreder, T., Kleemann, A. M., Schöpf, V., Kopietz, R., Anzinger, A., Demmel, M., Linn, J., Kettenmann, B., & Wiesmann, M. (2009). Olfactory detection thresholds and pleasantness of a food-related and a non-food odour in hunger and satiety. *Rhinology*, 47(2), 160–165.
- Auffarth, B. (2013). Understanding smell—The olfactory stimulus problem. *Neuroscience and Biobehavioral Reviews*, 37(8), 1667–1679. <https://doi.org/10.1016/j.neubiorev.2013.06.009>
- Boisson, C. (1997). La dénomination des odeurs: Variations et régularités linguistiques. *Intellectica*, 24(1), 29–49. <https://doi.org/10.3406/intel.1997.1546>
- Brand, G., & Millot, J.-L. (2001). Sex Differences in Human Olfaction: Between Evidence and Enigma. *The Quarterly Journal of Experimental Psychology Section B*, 54(3b), 259–270. <https://doi.org/10.1080/713932757>
- Bushdid, C., Magnasco, M. O., Vosshall, L. B., & Keller, A. (2014). Humans can Discriminate more than one Trillion Olfactory Stimuli. *Science (New York, N.Y.)*, 343(6177), 1370–1372. <https://doi.org/10.1126/science.1249168>
- Cain, W. S. (1979). To know with the nose: Keys to odor identification. *Science (New York, N.Y.)*, 203(4379), 467–470. <https://doi.org/10.1126/science.760202>
- David, S. (2002). Linguistic expressions for odors in French. *Olfaction, Taste and Cognition*, 82–99.
- Distel, H., Ayabe-Kanamura, S., Martínez-Gómez, M., Schicker, I., Kobayakawa, T., Saito, S., & Hudson, R. (1999). Perception of everyday odors—Correlation between intensity, familiarity and strength of hedonic judgement. *Chemical Senses*, 24(2), 191–199. <https://doi.org/10.1093/chemse/24.2.191>
- Doty, R. L. (2018). Age-Related Deficits in Taste and Smell. *Otolaryngologic Clinics of North America*, 51(4), 815–825. <https://doi.org/10.1016/j.otc.2018.03.014>
- Doty, R. L., & Kamath, V. (2014). The influences of age on olfaction: A review. *Frontiers in Psychology*, 5, 20. <https://doi.org/10.3389/fpsyg.2014.00020>

- Dravnieks, A. (1985). Atlas of odor character profiles. ASTM.
- Dubois, D., & Rouby, C. (2002). Names and categories for odors: The veridical label. *Olfaction, Taste and Cognition*, 10, 47–66.
- Ferdenzi, C., Roberts, S. C., Schirmer, A., Delplanque, S., Cekic, S., Porcherot, C., Cayeux, I., Sander, D., & Grandjean, D. (2013). Variability of affective responses to odors: Culture, gender, and olfactory knowledge. *Chemical Senses*, 38(2), 175–186. <https://doi.org/10.1093/chemse/bjs083>
- Jaeger, S. R., McRae, J. F., Bava, C. M., Beresford, M. K., Hunter, D., Jia, Y., Chheang, S. L., Jin, D., Peng, M., Gamble, J. C., Atkinson, K. R., Axten, L. G., Paisley, A. G., Tooman, L., Pineau, B., Rouse, S. A., & Newcomb, R. D. (2013). A Mendelian Trait for Olfactory Sensitivity Affects Odor Experience and Food Selection. *Current Biology*, 23(16), 1601–1605. <https://doi.org/10.1016/j.cub.2013.07.030>
- Keller, A., Gerkin, R. C., Guan, Y., Dhurandhar, A., Turu, G., Szalai, B., Mainland, J. D., Ihara, Y., Yu, C. W., Wolfinger, R., Vens, C., Schietgat, L., Grave, K. D., Norel, R., Consortium, D. O. P., Stolovitzky, G., Cecchi, G. A., Vosshall, L. B., & Meyer, P. (2017). Predicting human olfactory perception from chemical features of odor molecules. *Science*, 355(6327), 820–826. <https://doi.org/10.1126/science.aal2014>
- Keller, A., & Vosshall, L. B. (2004). Human olfactory psychophysics. *Current Biology*, 14(20), R875–R878.
- Keller, A., Zhuang, H., Chi, Q., Vosshall, L. B., & Matsunami, H. (2007). Genetic variation in a human odorant receptor alters odour perception. *Nature*, 449(7161), 468–472. <https://doi.org/10.1038/nature06162>
- Khan, R. M., Luk, C.-H., Flinker, A., Aggarwal, A., Lapid, H., Haddad, R., & Sobel, N. (2007). Predicting odor pleasantness from odorant structure: Pleasantness as a reflection of the physical world. *The Journal of Neuroscience: The Official Journal of the Society for Neuroscience*, 27(37), 10015–10023. <https://doi.org/10.1523/JNEUROSCI.1158-07.2007>
- Larsson, M., Öberg-Blåvarg, C., & Jönsson, F. U. (2009). Bad odors stick better than good ones: Olfactory qualities and odor recognition. *Experimental Psychology*, 56(6), 375–380. <https://doi.org/10.1027/1618-3169.56.6.375>
- Larsson, M., & Willander, J. (2009). Autobiographical odor memory. *Annals of the New York Academy of Sciences*, 1170(1), 318–323.
- Licon, C. C., Bosc, G., Sabri, M., Mantel, M., Fournel, A., Bushdid, C., Golebiowski, J., Robardet, C., Plantevit, M., Kaytoue, M., & Bensafi, M. (2019). Chemical features mining provides new descriptive structure-odor relationships. *PLoS Computational Biology*, 15(4), e1006945. <https://doi.org/10.1371/journal.pcbi.1006945>
- Licon, C. C., Manesse, C., Dantec, M., Fournel, A., & Bensafi, M. (2018). Pleasantness and trigeminal sensations as salient dimensions in organizing the semantic and physiological spaces of odors. *Scientific Reports*, 8(1), 8444. <https://doi.org/10.1038/s41598-018-26510-5>

- Mainland, J. D., Keller, A., Li, Y. R., Zhou, T., Trimmer, C., Snyder, L. L., Moberly, A. H., Adipietro, K. A., Liu, W. L. L., & Zhuang, H. (2014). The missense of smell: Functional variability in the human odorant receptor repertoire. *Nature Neuroscience*, 17(1), 114.
- Majid, A. (2015). Cultural Factors Shape Olfactory Language. *Trends in Cognitive Sciences*, 19(11), 629–630. <https://doi.org/10.1016/j.tics.2015.06.009>
- Majid, A., & Burenhult, N. (2014). Odors are expressible in language, as long as you speak the right language. *Cognition*, 130(2), 266–270. <https://doi.org/10.1016/j.cognition.2013.11.004>
- Majid, A., & Kruspe, N. (2018). Hunter-Gatherer Olfaction Is Special. *Current Biology*, 28(3), 409–413.e2. <https://doi.org/10.1016/j.cub.2017.12.014>
- Mantel, M., Ferdenzi, C., Roy, J.-M., & Bensafi, M. (2019). Individual Differences as a Key Factor to Uncover the Neural Underpinnings of Hedonic and Social Functions of Human Olfaction: Current Findings from PET and fMRI Studies and Future Considerations. *Brain Topography*, 32(6), 977–986. <https://doi.org/10.1007/s10548-019-00733-9>
- Mathonat, R., Nurbakova, D., Boulicaut, J.-F., & Kaytoue, M. (2021). Anytime mining of sequential discriminative patterns in labeled sequences. *Knowledge and Information Systems*, 63(2), 439–476. <https://doi.org/10.1007/s10115-020-01523-7>
- McGann, J. P. (2017). Poor human olfaction is a 19th-century myth. *Science (New York, N.Y.)*, 356(6338). <https://doi.org/10.1126/science.aam7263>
- Micoulaud-Franchi, J.-A., Balzani, C., & Vion-Dury, J. (2012). The explicitation interview in medicine: Exploring the subjective experiences. *Médecine*, 8(8), 363–367. <https://doi.org/10.1684/med.2012.0877>
- Obrist, M., Seah, S. A., & Subramanian, S. (2013). Talking about tactile experiences. *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, 1659–1668. <https://doi.org/10.1145/2470654.2466220>
- Pasquier, N., Bastide, Y., Taouil, R., & Lakhal, L. (1999). Discovering Frequent Closed Itemsets for Association Rules. *Proceedings of the 7th International Conference on Database Theory*, 398–416.
- Petitmengin, C. (2006). Describing one's subjective experience in the second person: An interview method for the science of consciousness. *Phenomenology and the Cognitive Sciences*, 5(3), 229–269. <https://doi.org/10.1007/s11097-006-9022-2>
- Petitmengin, C., van Beek, M., Bitbol, M., Nissou, J.-M., & Roepstorff, A. (2019). Studying the experience of meditation through Micro-phenomenology. *Current Opinion in Psychology*, 28, 54–59. <https://doi.org/10.1016/j.copsyc.2018.10.009>
- Rouby, C., Pouliot, S., & Bensafi, M. (2009). Odor hedonics and their modulators. *Food Quality and Preference*, 20(8), 545–549. <https://doi.org/10.1016/j.foodqual.2009.05.004>
- Royet, J.-P., Plailly, J., Saive, A.-L., Veyrac, A., & Delon-Martin, C. (2013). The impact of expertise in olfaction. *Frontiers in Psychology*, 4. <https://doi.org/10.3389/fpsyg.2013.00928>

Sezille, C., Fournel, A., Rouby, C., Rinck, F., & Bensafi, M. (2014). Hedonic appreciation and verbal description of pleasant and unpleasant odors in untrained, trainee cooks, flavorists, and perfumers. *Frontiers in Psychology*, 5, 12. <https://doi.org/10.3389/fpsyg.2014.00012>

Sorokowski, P., Karwowski, M., Misiak, M., Marczak, M. K., Dziekan, M., Hummel, T., & Sorokowska, A. (2019). Sex Differences in Human Olfaction: A Meta-Analysis. *Frontiers in Psychology*, 10. <https://doi.org/10.3389/fpsyg.2019.00242>

Stevenson, R. J. (2010). An Initial Evaluation of the Functions of Human Olfaction. *Chemical Senses*, 35(1), 3–20. <https://doi.org/10.1093/chemse/bjp083>

Wise, P. M., Olsson, M. J., & Cain, W. S. (2000). Quantification of Odor Quality. *Chemical Senses*, 25(4), 429–443. <https://doi.org/10.1093/chemse/25.4.429>

Annexe 2 : « Perte de l'odorat avec la COVID-19 et impact sur la qualité de vie » (publié)



Original Article

Recovery From COVID-19-Related Olfactory Disorders and Quality of Life: Insights From an Observational Online Study

Camille Ferdenzi^o, Christophe Bousquet^o, Pierre-Emmanuel Aguera, Morgane Dantec, Christelle Daudé, Lesly Fornoni, Arnaud Fournel, Aurélien Kassan, Marylou Mantel^o, Maëlle Moranges, Erwan Moussy, Stéphane Richard Ortegón, Catherine Rouby^o and Moustafa Bensafi^o

Lyon Neuroscience Research Center, CNRS UMR5292, INSERM U1028, University Claude Bernard Lyon 1, CH Le Vinatier, Bât. 462 Neurocampus, 95 boulevard Pinel, 69675 Bron Cedex, France

Correspondence to be sent to: Camille Ferdenzi, Lyon Neuroscience Research Center, CNRS UMR5292, INSERM U1028, University Claude Bernard Lyon 1, CH Le Vinatier, Bât. 462 Neurocampus, 95 boulevard Pinel, 69675 Bron Cedex, France.
e-mail: camille.ferdenzi@cnrs.fr

Editorial Decision 31 May 2021.

Abstract

Although olfactory disorders (OD) are among the most significant symptoms of COVID-19, recovery time from COVID-19-related OD and their consequences on the quality of life remain poorly documented. We investigated the characteristics and behavioral consequences of COVID-19-related OD using a large-scale study involving 3111 French respondents (78% women) to an online questionnaire over a period of 9 months covering different epidemic waves (from 8 April 2020 to 13 January 2021). In the patients who subjectively recovered from COVID-19-related OD ($N = 609$), recovery occurred on average after 16 days and most of the time within 1 month (“normal” recovery range); 49 subjectively recovered in 1–2.5 months, and several cases took up to 6.5 months. Among the patients with ongoing OD ($N = 2502$), 974 were outside the “normal” recovery range (persistent OD) and reported OD for 1–10 months. Developing a persistent OD was more likely with increasing age and in women and was more often associated with parosmia and phantosmia. The deleterious impact of COVID-19-related OD on the quality of life was significantly aggravated by OD duration and was more pronounced in women. Because persistent OD is not infrequent after COVID-19, has deleterious consequences on the quality of life, and receives few solutions from the health practitioners, it would be beneficial to implement screening and treatment programs to minimize the long-term behavioral consequences of COVID-19-related OD.

Key words: COVID-19, mental health, quality of life, self-report, smell loss

Introduction

Disturbances of smell and taste perception are among the most significant symptoms of the recently emerged COVID-19 disease. Several studies based on self-reports recorded a relatively high prevalence in many countries: 46–47% of COVID-19 patients were

found to have olfactory disorders (OD) in France (Bénézit et al., 2020; Klopfenstein et al., 2020), 32% in Spain (Beltrán-Corbellini et al., 2020), 19–24% in Italy (Giacomelli et al., 2020; Vaira et al., 2020), 49–65% in Germany (Haehner et al., 2020; Hornuss et al., 2020), 59% in the United Kingdom (Menni et al., 2020), 59% in the

United States (Yan et al., 2020), 28% in Iran (Moein et al., 2020), and 36% in Israël (Levinson et al., 2020). This prevalence was significantly higher in patients than in COVID-19-negative controls (Beltrán-Corbellini et al., 2020; Bénézit et al., 2020; Hornuss et al., 2020; Menni et al., 2020; Moein et al., 2020). The actual prevalence may even be higher because studies relying on subjective assessments of OD typically produce lower estimates of prevalence than studies relying on objective assessments (Desiato et al., 2021).

So far, the characteristics of COVID-19-related OD have been described as follows. They appear suddenly rather than gradually (Bagheri et al., 2020), total loss of olfactory perceptions (anosmia) is more frequent than hyposmia (decreased olfactory perceptions) (Beltrán-Corbellini et al., 2020; Lechien et al., 2020; Moein et al., 2020; but see Giacomelli et al., 2020; Hornuss et al., 2020), and OD are often concomitant with taste disorders (Bénézit et al., 2020; Giacomelli et al., 2020; Levinson et al., 2020; Moein et al., 2020; Parma et al., 2020). Investigation of gender and age differences yielded heterogeneous conclusions, with female patients being more frequently affected by OD in some studies (Giacomelli et al., 2020; Haehner et al., 2020; Lechien et al., 2020), but not in others (Beltrán-Corbellini et al., 2020; Moein et al., 2020), and patients with OD being younger than those without (Beltrán-Corbellini et al., 2020; Giacomelli et al., 2020), which is however not confirmed in other studies involving younger samples (Haehner et al., 2020; Hornuss et al., 2020). Finally, although first reports indicated that COVID-19-related OD last about 7–9 days on average (Beltrán-Corbellini et al., 2020; Klopfenstein et al., 2020; Lechien et al., 2020; Levinson et al., 2020), more recent evidence shows that 37% of the patients report persistent OD at 1.5 month after the first consultation and 14% report a partial (not total) recovery (Chiesa-Estomba et al., 2020). At 6-month follow-up, 5–11% of patients were found to still suffer from OD (Huang et al., 2021; Lechien, Chiesa-Estomba, Beckers et al., 2021).

Past studies performed before the COVID-19 pandemics have shown that OD can have serious consequences on quality of life (Vennemann et al., 2008; Keller and Malaspina, 2013; Croy et al., 2014; Philpott and Boak, 2014; Manesse et al., 2017; Drareni et al., 2019). Patients with impaired sense of smell are more likely to experience domestic accidents, report feelings of insecurity in their social relationships, as well as report alterations of eating behavior, which can lead to a tendency to isolation and depression (Croy et al., 2014; Philpott and Boak, 2014). Furthermore, women (Philpott and Boak, 2014) and current smokers (Vennemann et al., 2008) seem to have more severe impacts of OD on their quality of life.

How COVID-19 patients with OD recover their perceptual abilities on the longer term, and how this relates to individual characteristics, are currently poorly documented. The first aim of this observational study was to describe the various forms of OD reported by the participants, using online self-reports of patients who declared having been diagnosed with COVID-19 and having lost their sense of smell between February 2020 and January 2021. The second aim was to investigate the temporal dynamics of subjective recovery from OD due to COVID-19, and the factors involved in the cases of slower recovery. Persistent OD have been examined more thoroughly, given their anticipated deleterious consequences for the patients. The third aim was to investigate the consequences of COVID-19-related OD on patients' behavior, by linking the characteristics of a participant and of his/her OD to the quality of life assessed with a 16-item questionnaire. Finally, the fourth aim of this study was to describe how ODs were managed by the medical staff in the context of the COVID-19 pandemic. The impact of

COVID-19-related OD on quality of life as a function of recovery dynamics is an important societal concern, which has clinical implications regarding whether and how these sensory disorders require medical care.

Materials and methods

Participants

The online questionnaire was advertised at the national level in France and intended for people with OD, be it caused by COVID-19 or not. In this article, we included participants who 1) reported OD and 2) were diagnosed COVID-19 positive by a health practitioner from 1 February 2020 onwards and based on a laboratory test (74%: PCR test after nasal swab 62%, chest radio 0.2%, other, e.g., blood test 11.4%, unknown 0.4%) or based on their symptoms (26%). Although the diagnosis based only on the symptoms has a higher rate of uncertainty, we used it as an inclusion criterion because at the beginning of the pandemic many patients were diagnosed this way due to very limited access to PCR tests. Note that the conclusions of this article remain the same if only participants reporting a PCR test are included. Other inclusion criteria were as follows: 3) responding to the full questionnaire (i.e., no drop out before the end of the questionnaire), 4) completing the questionnaire for the first time, 5) aged over 18 years old, 6) answering between 8 April 2020 (questionnaire release) and 13 January 2021, and 7) providing usable data on OD duration and declaring OD onset maximum 1 week prior to COVID-19 onset. In total, responses of 3111 individuals were used. Participants were mostly women (78.1%) and were aged 40.5 ± 12.5 years (range 18–85). Men and women did not differ on age (mean \pm SD: 40.2 ± 13.1 and 40.6 ± 12.4 years, respectively; $t_{3109} = 0.751$, $P = 0.453$). The study was approved by the ethics committee of the Institute of Biological Sciences of the CNRS on 3 April 2020. All individuals provided informed consent when participating in the survey.

Questionnaire

The online questionnaire was accessible from a French information website dedicated to the sense of smell (<https://project.crnlf.fr/odorat-info/>). It was advertised at the national level in France through different channels (numerous audio-visual media communications, authors' professional and personal networks, advertisements posted in local health centers and pharmacies). The call was directed to people who noticed a change in their sense of smell, be it caused by COVID-19 or not, and be it still present or not. The stated purpose of the study was to help us better understand how OD manifest themselves and what impact they have on quality of life, especially in the context of the COVID-19 crisis. The questionnaire comprised 4 sections (see details in [Supplementary Appendix](#) and in Data Analysis section for the relevant variables we used in this article): 1) chemosensory disorders (smell, taste, trigeminal); characteristics and management by health practitioners, 2) incapacitating character of chemosensory disorders and impact on quality of life (QoL) on 16 items related to food, social interactions, danger detection and enjoyment of the environment, 3) sociodemographic information, and 4) information related with COVID-19. The questionnaire was completed only once by the participants (not meant for a follow-up), providing information at a single timepoint. Participants responded at varied times after the onset of the disease (and after the onset of their COVID-19-related OD), as illustrated in [Supplementary Figure S1](#).

Data analysis

Variables

For data analysis, we calculated the duration of a participant's OD, based on the reported date of OD onset and the reported date of OD disappearance, for individuals who subjectively recovered. Age, gender, height and weight (from which we computed body mass index [BMI]), smoking status (yes or no), type of OD (total, i.e., anosmia, or partial, i.e., hyposmia), speed of OD onset (sudden or progressive), and constancy of OD (fluctuating or constant) were also included. For the quality of life, 3 variables were used. First, the response to the question whether OD was incapacitating for a participant's personal life (yes/no). Second, the 16 QoL item scores (−1, 0, 1 corresponding to the answers “less than usual,” “neither more nor less than usual,” “more than usual”). Third, a variable called TotalImpact (ranging from 0 to 16) summarizing the 16 QoL items and which is equal to the number of times a respondent answered either “less than usual” or “more than usual” (i.e., deviated from “neither more nor less than usual”).

Analyses

Analyses were performed in R 4.0.3 (R Core Team, 2021) and the level of significance was set at $\alpha = 0.05$. To facilitate the interpretation of the regression coefficients, all numerical variables were centered and scaled prior to inclusion in all models (Schielzeth, 2010).

First, we aimed at identifying the factors associated with OD duration for participants who subjectively recovered from OD. We used a generalized linear model with the positive negative binomial family (vglm function from the R package VGAM, Yee, 2021) and participant's 1) age, 2) gender, 3) BMI, 4) smoking status, 5) OD type, 6) OD onset speed, 7) OD constancy, and 8) age by gender interaction as explanatory variables. Age by gender interaction was included because such interactions may occur in olfaction (e.g., Manesse et al., 2021). A backward model selection, based on likelihood ratio tests (LRTs) and using the lrttest function from the R package lmerTest (Zeileis and Hothorn, 2002), allowed us to obtain a minimal model containing only variables having a significant effect.

Second, we aimed at identifying the factors influencing the probability to report a persistent OD or to have subjectively recovered. We started by focusing on defining the “normal” duration range of recovery from OD in our sample. To do so, rules using the interquartile range (IQR) of the distribution are routinely used (Rousseeuw and Hubert, 2011). With this technique, points outside the range [first quartile − 1.5 × IQR, third quartile + 1.5 × IQR] are considered as outliers. In our case, the “normal” range is thus [7 − 1.5 × 10, 17 + 1.5 × 10] or [−8, 32] and OD durations longer than 32 days are considered as outliers. Then, we filtered out respondents with ongoing OD but who were still within the “normal” range (1–32 days) because our approach does not allow us to know whether in the future these participants will be recovering within the “normal” range or whether they will develop longer-lasting OD. We defined a new independent binary variable—persistence of OD—which is equal to 1 when respondents had persistent OD (>32 days, not recovered yet, $N = 974$) and 0 when respondents had subjectively recovered from their OD ($N = 609$). This variable was modeled by a logistic regression with the same 8 explanatory variables described above. Again, model selection allowed us to only retain the significant explanatory variables (stepwise model selection based on Akaike information criteria (AIC) with the step function from the R package stats).

Third, we aimed at identifying the variables that predicted the impacts of OD on the quality of life of all participants ($N = 3111$). We ran a linear regression on the composite variable TotalImpact with 9 explanatory variables, which are the same 8 already described

to which we added (9) the duration of OD, and a stepwise model selection based on AIC was conducted. Then, we analyzed the 16 QoL item scores. To determine whether the impact was significant, we compared the distribution of the answers “neither more nor less than usual” versus “less than usual” and “more than usual” grouped together, using raw numbers converted in percentage rounded to the nearest whole number, with the theoretical distribution if OD induced no change (i.e., 100/0), using χ^2 tests. To describe the direction of the impact (reduction or increase), we compared with χ^2 tests the number of respondents who answered “More than usual” to the number of “Less than usual” responses (irrespective of how many respondents answered “Same as usual”). P -values were Bonferroni corrected for repeated testing (significant at $\alpha = 0.05/16$ items = 0.003125). To determine the variables explaining the variations in responses to each quality of life item, we performed multivariate ordinal regressions (MORs) using the mvord function in the R package mvord (Hirk et al., 2020), including the same 9 explanatory variables described just before. Due to computation time constraints, it was not possible to run a single MOR for the 16 QoL items: 4 MORs were therefore run on 4 subgroups of items that were identified with a factor analysis (see details in Supplementary Method and Supplementary Figure S2). Again, model selection using LRT was performed to only retain the significant explanatory variables. As 4 MORs were conducted instead of 1, a Bonferroni correction was applied to the P -values, reported as significant at $\alpha = 0.05/4 = 0.0125$.

Results

Self-reported characteristics of COVID-19-related ODs

COVID-19-related ODs were often associated with taste disorders (71.0% of the 3111 included respondents) and in 25.5% of the cases with a dysfunction of the trigeminal sensitivity. Only 15.0% of the COVID-19 patients with OD reported having a blocked nose when they underwent the dysfunction. All included participant had decreased olfactory perceptions (quantitative OD), but qualitative disorders were also described: parosmia (odors smelling different from what they usually do) in 55.8% of the respondents, and phantosmia (olfactory hallucinations, see detailed descriptions in Supplementary Results) in 34.6%. Note that 26.6% reported having both parosmia and phantosmia. The quantitative alterations were characterized by 1) type of OD: a clear predominance of anosmia (86.4% vs. 13.6% hyposmia), 2) onset speed of OD: a sudden onset (86.8% vs. 13.2% progressive onset), and 3) constancy of OD: a constant alteration (67.4% vs. 32.6% fluctuating). How the frequency of these characteristics were affected by age and gender was analyzed using logistic regressions: Women had a higher probability than men to report a fluctuating rather than constant OD ($\beta = 0.22 \pm 0.10$, $z = 2.36$, $P = 0.02$; no gender differences for type or speed of OD: $\beta = 0.02 \pm 0.13$, $z = 0.15$, $P = 0.881$ and $\beta = 0.24 \pm 0.14$, $z = 1.76$, $P = 0.08$, respectively). Moreover, increasing age was associated with an increased probability to report a sudden rather than progressive onset ($\beta = 0.29 \pm 0.06$, $z = 5.12$, $P < 0.001$) and a fluctuating rather than constant OD ($\beta = 0.14 \pm 0.04$, $z = 3.66$, $P < 0.001$) (no effect of age on OD type: $\beta = 0.003 \pm 0.052$, $z = 0.06$, $P = 0.955$).

Self-reported timing of recovery from COVID-19-related OD

Participants who subjectively recovered from OD

Among the 3111 COVID-19 patients with OD who completed the online questionnaire, 609 (19.6%) reported that they had recovered

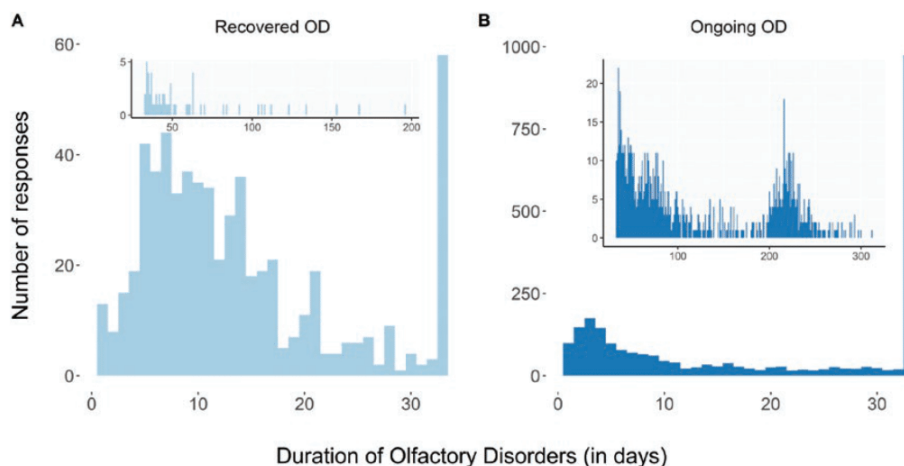


Figure 1. (A) Distribution of the self-reported OD duration by participants who recovered their sense of smell ($N = 609$, 90.5% recovered within 32 days or less). (B) Distribution of the self-reported OD duration by the participants who still suffer from OD ($N = 2502$; 974 participants have persistent OD, i.e. they have not recovered after 32 days). The insets display the distributions of OD durations longer than 32 days. On the inset of (B), the distribution between 200 and 300 days corresponds to respondents who contracted COVID-19 during the first wave of the disease in spring 2020 ($N = 329$). Note the different scales on the x and y axes between the main panels and the insets.

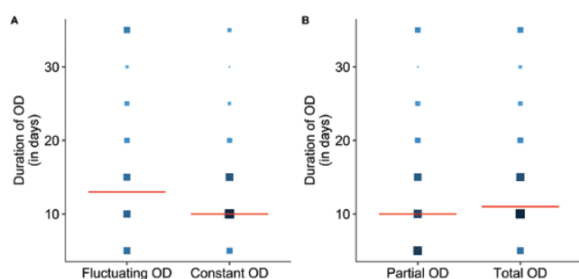


Figure 2. Effects of (A) constancy of OD on OD duration (longer duration for participants reporting fluctuating OD) and (B) type of OD on OD duration (longer duration for participants reporting total OD) ($N = 609$). The size and the color darkness of the squares are proportional to the number of participants reporting fluctuating, constant, partial, or total OD; the horizontal red lines represent the median OD duration (in days). For illustration purposes, the days have been grouped by 5 (however, the median and the analysis take the whole distribution of durations into account).

their smell. In these patients, subjective recovery took 16.3 ± 19.3 days on average (median: 11 days), ranging from 1 (appears and disappears the same day) to 196 days (Figure 1A). This recovery duration was significantly predicted by constancy and type of OD (but not gender, age, BMI, smoking status, or onset speed of OD). Patients with fluctuating (vs. constant) OD (Figure 2A) and with a total (vs. partial) OD (Figure 2B) took longer to subjectively recover from their OD (zero-truncated negative binomial minimal model: $\beta = 0.26 \pm 0.07$, $z = 3.65$, $P < 0.001$ and $\beta = 0.33 \pm 0.14$, $z = 2.30$, $P < 0.05$, respectively).

Persistent OD

Among the COVID-19 patients with OD who had not subjectively recovered their smell at the time they completed the questionnaire ($N = 2502$), 1528 were within what could be called the “normal” range of recovery time, that is, from 1 to 32 days (see Materials and methods). Nine hundred and seventy-four were outside this range and thus considered as displaying persistent COVID-19-related OD (Figure 1B). These patients represent almost one third (31.3%)

of the 3111 participants spontaneously responding to our online questionnaire. They reported OD durations of up to 312 days (i.e., ~10 months). Four factors significantly differentiated this subgroup (persistent OD) from the group who subjectively recovered (Figure 3A). The participants with persistent OD were much more likely to have partial compared to total olfactory loss (logistic regression minimal model: $\beta = 1.63 \pm 0.18$, $z = 8.82$, $P < 2e-16$, OR = 5.01 [3.59–7.41], Figure 3B). Women were more likely to report persistent OD compared to men (logistic regression minimal model: $\beta = 0.65 \pm 0.13$, $z = 5.10$, $P < 0.001$, OR = 1.92 [1.50–2.48], Figure 3C). Furthermore, persistent OD were more likely to have a sudden onset rather than a progressive one (logistic regression minimal model: $\beta = 0.61 \pm 0.17$, $z = 3.50$, $P < 0.001$, OR = 1.84 [1.31–2.60], Figure 3D). Finally, older respondents reported persistent OD more frequently (logistic regression minimal model: $\beta = 0.32 \pm 0.05$, $z = 5.75$, $P < 0.001$, OR of each additional year = 1.38 [1.24–1.54], Figure 3E). The constancy of OD, smoking status, BMI, and the interaction between gender and age were not retained in the minimal model. Finally, participants with persistent OD were much more likely to declare a parosmia or a phantosmia ($\chi^2 = 68.1$, $df = 1$, $P < 0.001$ and $\chi^2 = 67.6$, $df = 1$, $P < 0.001$, respectively, χ^2 test comparing patients with persistent OD vs. patients who subjectively recovered; Figure 4A and B).

COVID-19-related OD and self-reported quality of life

Disabling character

When asked whether losing their smell was incapacitating for their personal life, 73.1% ($N = 2275$ of 3111) of the COVID-19 patients with OD considered that it was.

QoL questionnaire: total impact and factors of variation.

The total impact was significantly greater than 0 (intercept in the minimal model: $\beta = 4.35 \pm 0.18$, $t = 23.85$, $P < 2e-16$). Six factors were found to significantly influence the TotalImpact of OD on a respondent's quality of life (i.e., number of situations where the respondents reported behaving differently than usual). OD disturbed more facets of quality of life when it lasted longer (minimal model: $\beta = 0.67 \pm 0.06$, $t = 12.14$, $P < 2e-16$, Figure 5A). Younger respondents

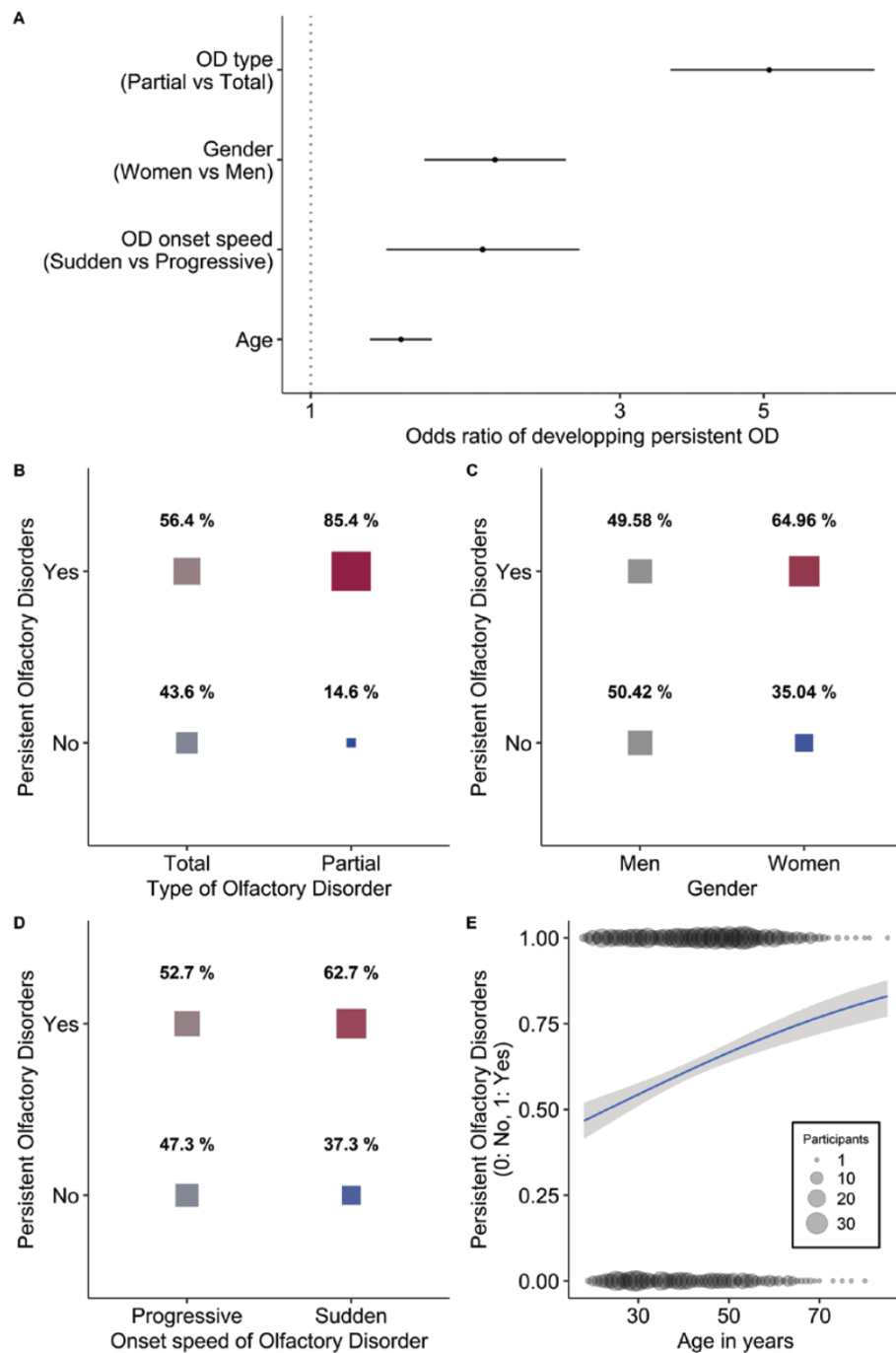


Figure 3. Four factors linked to the development of persistent OD (versus recovered). (A) Odds ratios of the 4 factors (of 8) that were retained in the minimal model of the logistic regression explaining persistent OD (vs. recovered); logarithmic scale. Percentage of cases of persistent OD versus recovered as a function of (B) OD type (total, partial), (C) gender (men, women), (D) OD onset speed (progressive, sudden), and (E) age. The color and size of squares are proportional to the reported proportions. The blue line and gray-shaded area correspond to the predicted values and their 95% confidence interval.

reported higher total impact than older respondents did (minimal model: $\beta = -0.45 \pm 0.05$, $t = -8.58$, $P < 2e-16$, Figure 5B). Women reported a broader impact of OD on their quality of life than men (minimal model: $\beta = 1.48 \pm 0.12$, $t = 11.96$, $P < 2e-16$, Figure 5C). Total OD had stronger effects on quality of life than partial OD (minimal model: $\beta = 1.27 \pm 0.15$, $t = 8.22$, $P = 3e-16$, Figure 5D). Similarly,

constant OD had stronger effects on quality of life than fluctuating OD (minimal model: $\beta = 0.50 \pm 0.12$, $t = 4.28$, $P = 2e-05$, Figure 5E). Finally, smokers reported a stronger impact of OD on their quality of life than nonsmokers (minimal model: $\beta = 0.44 \pm 0.13$, $t = 3.40$, $P = 0.0007$, Figure 5F). Onset speed of OD, BMI, and the interaction between gender and age were not retained in the minimal model.

QoL questionnaire: impact per item and factors of variation.

Considering the 16 quality of life items separately, the distribution of the answers (“neither more nor less than usual” versus “less than usual” + “more than usual”) significantly differed from theoretical distribution (no change: 100/0) for all items (all P s < 0.001) except Burn clothes ($P = 0.1552$), indicating significant impact of OD for 15 of 16 items. To describe the main direction of these changes, we compared the frequency of displaying each behavior “more” versus “less” than before losing the sense of smell (see [Supplementary](#)

[Figure S3](#)). Participants more frequently reported an increase than a decrease for 11 items (Seek close ones’ odor, Control own body odor, Add salt, Burn meals, Add spices, Prefer to eat alone, Eat rotten food, Take showers, Have domestic accidents, Add sugar and Add fat; χ^2 tests, all P s < 0.001, items in green in [Figure 6](#)). For the remaining 5 items, participants more frequently reported a decrease than an increase (Wear perfume, Detect smokes/gas, Enjoy smelling flowers, Enjoy eating; χ^2 -tests, all P s < 0.001, items in blue in [Figure 6](#)) or no difference (Burn clothes, $P = 0.018$ which is higher than the Bonferroni corrected alpha, item in black in [Figure 6](#)).

Among the most salient results obtained when considering each item of the quality of life questionnaire separately ([Supplementary Figures S4–S7](#)), we found that OD duration as well as gender and age had significant effects on many items (see summary of the effects in [Figure 6](#)). Specifically, a long-lasting OD increased the probability to report 1) adding salt, sugar, spices and fat to dishes, 2) experiencing domestic accidents, burning meals and eating rotten food, and 3) controlling one’s body odor, taking showers and seeking the smell of loved ones, more than usual. In addition, women, compared to men, reported 1) adding salt, 2) having domestic accidents and burning meals, 3) controlling body odor of self and the loved ones, and taking showers, more than usual, but also 4) having less pleasure smelling flowers and wearing perfume less often. Finally, younger participants reported eating rotten food more often than older ones, and displayed more variation (i.e., answering either “more” or “less”) in salting habits and eating pleasure. Results involving the other explanatory variables (age by gender interaction, BMI, smoking status and OD characteristics) can be found in [Supplementary Results](#).

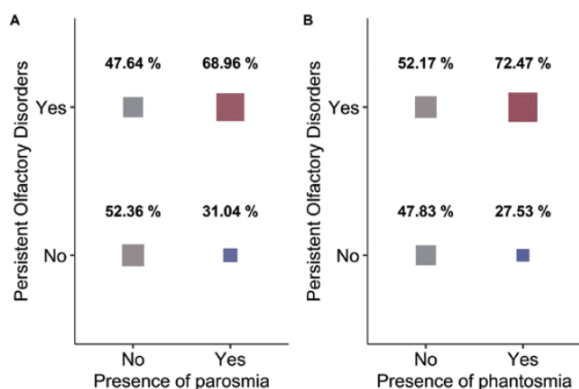


Figure 4. Frequency of (A) parosmia and (B) phantosmia in patients with persistent COVID-related OD (ongoing OD for more than 32 days, $N = 974$) and in patients who recovered from their COVID-related OD ($N = 609$). The color and size of squares are proportional to the reported proportions.

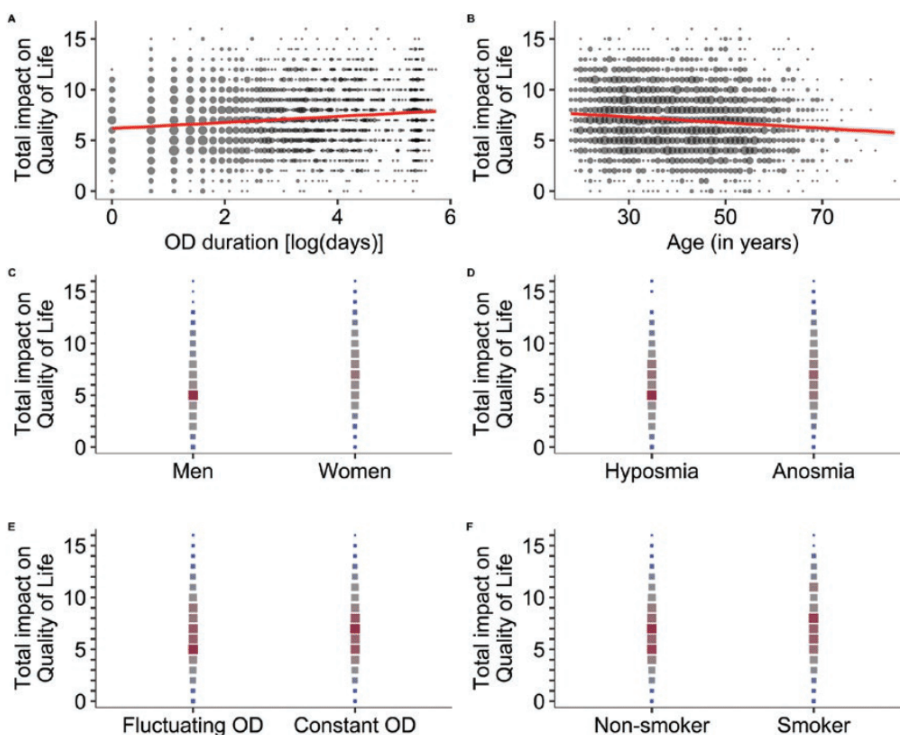


Figure 5. Six factors influencing significantly the impact of COVID-related OD on the patients’ quality of life: (A) OD duration, (B) age, (C) gender, (D) OD type, (E) OD frequency, and (F) smoking status ($N = 3111$). The blue line and gray-shaded area correspond to the predicted values and their 95% confidence interval. The color and size of squares are proportional to the number of participants reporting a specific number of impacts on quality of life.

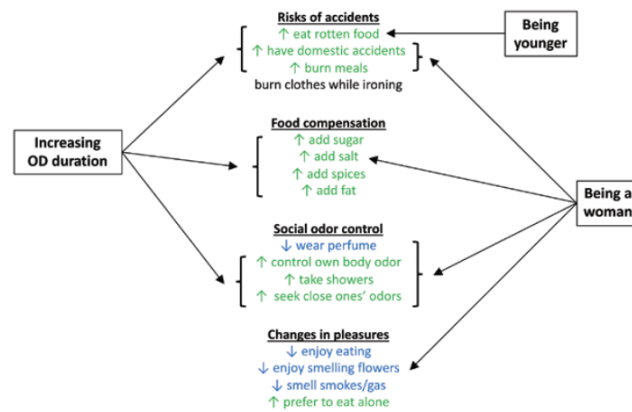


Figure 6. Effects of OD duration, gender, and age on each of the 16 items of the quality of life questionnaire ($N = 3111$). Items in green received more “more than usual” than “less than usual” answers, and items in blue had the reverse pattern (significant χ^2 tests; no difference for the item in black; see frequencies in Supplementary Figure S2).

Self-reported management of COVID-19-related OD

When questioned about the management of their OD in the context of COVID-19, less than half of the COVID-19 patients with OD (41.5%, $N = 1292$) reported that they consulted specifically for their sensory loss. Most of them (87.5%) were examined by a general practitioner (GP). Some patients (19.0%) consulted a specialist (ENT physician in 88.6% of the cases). In most cases, OD only served to orient the health practitioner toward a diagnosis of COVID-19 without considering them as a symptom to treat, despite the concern many patients had regarding their sensory loss. Treatment of olfactory loss was rarely evoked, but when it was, the practitioner's most frequent answers were either to just wait for a spontaneous recovery (18.3% of the consultations) or to try olfactory training (among other even more anecdotal options such as nasal sprays, nasal wash or vitamins). Olfactory training has been recommended by only 3.6% of the GP consultations and 18.4% of the ENT consultations, and took very heterogeneous forms (no standard protocol).

Discussion

Using self-reports of more than 3100 respondents collected over a period of 9 months covering different epidemic waves in France (from 8 April 2020 to 13 January 2021), this study showed that COVID-19-related OD 1) last longer than initially thought (several months) in many patients, 2) have significant impacts on the quality of life and that this impact is increasingly deleterious with OD duration, and 3) receive very few treatment options from the health practitioners despite their behavioral consequences.

From a methodological point of view, it must be kept in mind that this approach inevitably has limits compared with an experimental setting controlled by an experimenter. In particular, there might be some imprecision 1) in how participants self-evaluated their sense of smell (it is well known that we are poorly able to do so: Landis et al., 2003; Manesse et al., 2021), although we tried to minimize this imprecision by asking very specific and detailed questions (see Supplementary Appendix A), and 2) in how participants reported dates, especially for those who answered several months after the onset of their OD (see Supplementary Figure S1). For long-lasting OD, this questionnaire was not designed to allow for a systematic description of the variations of OD that could have occurred over time. Another source of imprecision is that we had to

trust participants with the positivity of their Covid test, or with their diagnosis when it was based on symptoms alone. There might be a selection bias such that the respondents are people willing to participate in scientific research and probably the most affected by their olfactory loss (as suggested also by the greater proportion of female participants, who are known to be more affected by smells and smell disorders; Martin et al., 2001; Nordin et al., 2004). Finally, this is not a follow-up study since the information is collected once at a single timepoint over the course of the disease, and no prevalence data can be extracted from this study. Bearing this in mind, we can however rely on the validity of our results, which are fully in line with previous studies: we found that COVID-19-related OD is characterized by a more frequent anosmia than hyposmia, with a sudden onset mostly, association with taste and trigeminal disorders, and with infrequent blocked nose as reported elsewhere (Lechien et al., 2020; Parma et al., 2020; Speth et al., 2020) and descriptions of the phantom smells are very similar to what has been found in other studies (burned, foul, rotten; Leopold, 2002). Moreover, the way we conducted our analyses allow us to draw valid conclusions in the sample we considered, especially by investigating the factors of variation of OD characteristics and OD consequences on the quality of life within the respondents.

Analysis of the COVID-19-related OD duration, limited to the patients who said that they recovered their sense of smell, indicated that subjective recovery occurred most of the time within a month (1–32 days for 90% of the healed patients) with an average of 16 days and a median of 11 days. The remaining 10% took more time to subjectively recover, mostly between 1 and 2.5 months, but sometimes more (up to 6 months). These numbers are slightly higher than those published earlier during the pandemic (7–9 days on average; Beltrán-Corbellini et al., 2020; Klopfenstein et al., 2020; Levinson et al., 2020), probably because these earlier studies did not have sufficient hindsight to examine longer healings. Indeed, more recent reports provide duration ranges that are consistent with our findings (most of the patients recovered by 30 days in D'Ascanio et al., 2021) and there is increasing evidence that part of the patients who lost their sense of smell following SARS-CoV-2 infection are slower to recover (Chiesa-Estomba et al., 2020; Huang et al., 2021). Our data sheds an additional light on those persistent ODs and reveals that many patients (one third of the respondents in our study) have not recovered within the first month, and report subjective OD durations up to 312 days (10 months). Although our study does not

allow us to establish the prevalence of these persistent ODs, this proportion is consistent with other reports on long-term sequelae (33–36% of incomplete recovery of olfactory and gustatory function, Willi et al., 2021). We noticed that several hundreds of respondents ($N = 329$) still suffered from OD since they contracted the virus during the first wave (spring 2020). Here, we must acknowledge that it is difficult to state with certainty that all patients who reported to be cured are really cured: Indeed, recurring testimony in open-ended questions of the survey revealed that patients can recover from total anosmia, spend several weeks with a seemingly normal sense of smell before experiencing qualitative disorders (parosmia, phantosmia).

Mechanistically, the discrepancies of recovery times (from a few days to several months) between individuals could be partly explained by how severely the SARS-CoV-2 virus has damaged the olfactory system (Cooper et al., 2020), even though the mechanisms are still discussed (Lechien, Chiesa-Estomba, Hans, et al. 2021). In the most benign cases, SARS-CoV-2 induces inflammatory processes in the nasal cavity that prevent odorants from reaching the olfactory epithelium (Eliezer et al., 2020). Concomitantly, olfactory neurons may be dysfunctional due to the local inflammation (Bryche et al., 2020). In such cases, recovery time should be less than 1 month (Eliezer et al., 2020; Bryche et al., 2020). In the most severe scenarios, cells of the olfactory epithelium are infected by the virus because they possess ACE2 receptors which are gateways to the cells. When sustentacular (supporting) cells are infected, structure damage and ionic imbalance cause the inactivation and eventually the death of the olfactory neurons (Cooper et al., 2020). When basal cells are infected, the consequences are likely to be even more severe because neural regeneration is heavily compromised. In the current state of knowledge, it seems that olfactory neurons themselves cannot be infected (Brann et al., 2020, but see Satarker and Nampoothiri, 2020; Meinhardt et al., 2021) although a very recent study contradicts this hypothesis (de Melo, 2021).

When investigating the parameters of the disease and the individual characteristics of the patients, we were able to identify which factors were the best predictors of the speed of subjective recovery. First, in the subjectively recovered group, total loss is associated with longer recovery, which could be due to a more severe damage/inflammation in the olfactory cleft and olfactory epithelium area. Longer subjective recovery times were also associated with more fluctuating OD. Second, the probability to develop a persistent OD (>32 days; compared with successful subjective recovery) was increased by increasing age and by gender, namely by being a woman. One can hypothesize that peripheral damage of the olfactory epithelium takes longer to spontaneously be repaired in the elderly, due to slower neural regeneration processes (Watabe-Rudolph et al., 2011), although this would need further testing. Regarding the gender differences, it cannot be excluded that the well-known higher awareness of women regarding smells in general make them better able to detect alterations of their perceptions, and to report them in such a study. Also, we found that partial OD and sudden onset were associated more with a persistent OD than with subjective recovery. Although sudden onset is difficult to interpret, partial OD is corroborated by the higher frequency of parosmia and phantosmia in persistent OD. Indeed, it seems that OD characteristics vary during the recovery process (which could unfortunately not be apprehended here since we did not repeat the measures for each respondent). Namely, after a period of severe loss of smell (quantitative loss), patients regain olfactory perceptions, but in some of them, these sensations are still not fully normal (qualitative alterations). Parosmia can be due to the fact that only some olfactory receptors are functional but not

all, thereby distorting the usual pattern of activation (Parker et al., 2021); Phantosmia is associated with disordered growth of olfactory axons (Leopold, 2002). Both mechanisms are likely to occur during recovery and neural regeneration.

The fact that 1) COVID-19-related OD can be particularly long-lasting, 2) it does not concern only several isolated cases but hundreds of people, and 3) we still do not have enough hindsight to formulate a prognosis for recovery/recovery time (which may therefore be underestimated in our study), should draw attention to the consequences for the patients' quality of life. Altogether, participants in our study very frequently reported that COVID-19-related OD are incapacitating (73%), and we found that their quality of life was significantly impacted in almost all tackled domains (food, social, alarm). Although we cannot fully exclude that it could have been worsened by the impact of the pandemic/lockdowns themselves since we did not compare with a control group, this is in accordance with previous studies on smell loss (Keller and Malaspina, 2013; Croy et al., 2014; Philpott and Boak, 2014; Manesse et al., 2017; Drareni et al., 2019). Most importantly, the impact on the quality of life is all the more so deleterious that duration of OD is long. More specifically, when OD is prolonged, food compensation by adding ingredients possibly harmful for nutritional balance (sugar, fat, salt) is increasing, therefore posing a threat to health. Longer ODs were also associated with a higher likelihood of domestic accidents (in general, but also ingesting spoiled food or letting a dish burn), which constitute a significant source of danger. Finally, behaviors related to own body odor and the odor of the loved ones were also increasingly disturbed with increasing OD duration. This probably reflects the decreasing ability to cope with the OD as it develops over time. Besides this significant modulation of the quality of life by OD duration, we found that age and gender were also significantly influential. Namely, women were more affected than men in several domestic and social contexts, in accordance with their higher attention to smells in such situations (Martin et al., 2001; Frasnelli and Hummel, 2005). Younger individuals' quality of life was more impacted than in the older ones: Although again we cannot fully exclude an effect of the pandemic/lockdowns, which could be more pronounced on the youngest, the decrease in olfactory capacities with age (Doty et al., 1984; Hummel et al., 1997) could also make the oldest less sensitive to the deleterious effects of olfactory alterations. Total and constant OD were found to be more deleterious on quality of life as well, which is understandable given the lack of respite associated with these forms of OD. Finally, smokers were more impacted as well, which is difficult to explain since they are known to have diminished olfactory functioning (Vennemann et al., 2008), but could be an indirect effect of more severe forms of the disease (as smoking leads to more severe forms of COVID-19: Engin et al., 2020; Li et al., 2021) rather than a direct effect of OD per se.

These deleterious consequences of OD for the COVID-19 patients' quality of life are important findings for the management of COVID-19-related smell disorders. This may affect the mental health of the patients by leading to mood disorders and depression (Kohli et al., 2016). In the current context, marked by the fear of the virus and of its new variants and by the feeling that this health crisis is unending, the management of mental health is crucial for public health (Fiorillo and Gorwood, 2020; Pfefferbaum and North, 2020). However, the management of OD by health practitioners remains clearly insufficient. Knowledge about smell disorders is poor in the general population (although it probably recently increased due to media coverage during the COVID-19 pandemic). In addition, smell loss is generally believed—also by medical staff—to be less disabling

than other sensory losses, and medical advice consequently is often insufficient (Landis et al., 2009). Our study confirmed this. When patients consulted about their olfactory loss, this symptom helped health practitioners to make the diagnosis of COVID-19 without being considered as a symptom to treat most of the time. A lack of knowledge about the treatment options was also noticed, since only 4% of general practitioners (18% of ENT specialists) recommended olfactory training, an interesting and easy-to-implement option to help recover olfaction (Hummel et al., 2009; Manesse et al., 2018).

To conclude, despite some limitations (possible aggravation of OD impact on quality of life by the pandemic/lockdowns themselves, women disproportionately represented in the studied samples, some degree of imprecision in the participants responses due to long delays between questionnaire completion and OD onset), our study provides pieces of evidence for the following. COVID-19-related smell disorders appear to be quite severe, can be reported as long-lasting, and seem to significantly degrade the patients' quality of life thereby constituting a potential threat to their mental and physical health. These disorders should thus be considered seriously by health practitioners and public health decision-makers in a context which is currently highly stressing and socially isolating. Careful screening of COVID-19-related smell disorders and the development of treatment programs are highly encouraged to minimize the long-term cognitive and behavioral consequences of COVID-19.

Supplementary material

Supplementary data are available at *Chemical Senses* online.

Funding

This work was carried out with the financial support of the IDEXLYON Project of the University of Lyon as part of the Future Investments Program (ANR-16-IDEX-0005, CORODORAT project to CF and MB), and the Fondation Fyssen Paris (grant no. 173867 to C.F.).

Acknowledgments

The authors thank all the people who helped disseminate the questionnaire.

References

- Bagheri SH, Asghari A, Farhadi M, Shamshiri AR, Kabir A, Kamrava SK, Jalessi M, Mohebbi A, Alizadeh R, Honarmand AA, et al. 2020. Coincidence of COVID-19 epidemic and olfactory dysfunction outbreak in Iran. *Med J Islam Repub Iran*. 34:62.
- Beltrán-Corbellini Á, Chico-García JL, Martínez-Poles J, Rodríguez-Jorge F, Natera-Villalba E, Gómez-Corral J, Gómez-López A, Monreal E, Parra-Díaz P, Cortés-Cuevas JL, et al. 2020. Acute-onset smell and taste disorders in the context of COVID-19: a pilot multicentre polymerase chain reaction based case-control study. *Eur J Neurol*. 27:1738–1741.
- Bénézit F, Le Turnier P, Declerck C, Paillé C, Revest M, Dubée V, Tattevin P, Arvieux C, Baldeyrou M, Chaplain J-M, et al. 2020. Utility of hyposmia and hypogeusia for the diagnosis of COVID-19. *Lancet Infect Dis*. 20:1014–1015.
- Brann DH, Tsukahara T, Weinreb C, Lipovsek M, den Berge KV, Gong B, Chance R, Macaulay IC, Chou H-J, Fletcher RB, et al. 2020. Non-neuronal expression of SARS-CoV-2 entry genes in the olfactory system suggests mechanisms underlying COVID-19-associated anosmia. *Sci Adv*. 6:eabc5801.
- Bryche B, St Albin A, Murri S, Lacôte S, Pulido C, Ar Gouilh M, Lesellier S, Servat A, Wasniewski M, Picard-Meyer E, et al. 2020. Massive transient damage of the olfactory epithelium associated with infection of sustentacular cells by SARS-CoV-2 in golden Syrian hamsters. *Brain Behav Immun*. 89:579–586.
- Chiesa-Estomba CM, Lechien JR, Radulesco T, Michel J, Sowerby LJ, Hopkins C, Saussez S. 2020. Patterns of smell recovery in 751 patients affected by the COVID-19 outbreak. *Eur J Neurol*. 27:2318–2321.
- Cooper KW, Brann DH, Farruggia MC, Bhutani S, Pellegrino R, Tsukahara T, Weinreb C, Joseph PV, Larson ED, Parma V, et al. 2020. COVID-19 and the chemical senses: supporting players take center stage. *Neuron*. 107(2):219–233.
- Croy I, Nordin S, Hummel T. 2014. Olfactory disorders and quality of life – an updated review. *Chem Senses*. 39(3):185–194.
- D'Ascanio L, Pandolfini M, Cingolani C, Latini G, Gradoni P, Capalbo M, Frausini G, Maranzano M, Brenner MJ, Di Stadio A. 2021. Olfactory dysfunction in COVID-19 patients: prevalence and prognosis for recovering sense of smell. *Otolaryngol Head Neck Surg*. 164(1):82–86.
- de Melo G, Lazarini F, Levallois S, Hautefort C, Michel V, Larrous F, Verillaud B, Aparicio C, Wagner S, Gheusi G, et al. 2021. COVID-19-related anosmia is associated with viral persistence and inflammation in human olfactory epithelium and brain infection in hamsters. *Sci Transl Med*. 13:eabf8396.
- Desiato VM, Levy DA, Byun YJ, Nguyen SA, Soler ZM, Schlosser RJ. 2021. The prevalence of olfactory dysfunction in the general population: a systematic review and meta-analysis. *Am J Rhinol Allergy*. 35(2):195–205.
- Doty RL, Shaman P, Applebaum SL, Giberson R, Sikorski L, Rosenberg L. 1984. Smell identification ability: changes with age. *Science*. 226(4681):1441–1443.
- Drareni K, Dougkas A, Giboreau A, Laville M, Souquet PJ, Bensafi M. 2019. Relationship between food behavior and taste and smell alterations in cancer patients undergoing chemotherapy: a structured review. *Semin Oncol*. 46(2):160–172.
- Eliezer M, Hamel A-L, Houdart E, Herman P, Housset J, Jourdaine C, Eloit C, Verillaud B, Hautefort C. 2020. Loss of smell in COVID-19 patients: MRI data reveals a transient edema of the olfactory clefts. *Neurology*. 95(23):e3145–e3152.
- Engin AB, Engin ED, Engin A. 2020. Two important controversial risk factors in SARS-CoV-2 infection: Obesity and smoking. *Environ Toxicol Pharmacol*. 78:103411.
- Fiorillo A, Gorwood P. 2020. The consequences of the COVID-19 pandemic on mental health and implications for clinical practice. *Eur Psychiatry*. 63(1):e32.
- Frasnelli J, Hummel T. 2005. Olfactory dysfunction and daily life. *Eur Arch Otorhinolaryngol*. 262(3):231–235.
- Giacomelli A, Pezzati L, Conti F, Bernacchia D, Siano M, Oreni L, Rusconi S, Gervasoni C, Ridolfo AL, Rizzardini G, et al. 2020. Self-reported olfactory and taste disorders in SARS-CoV-2 patients: a cross-sectional study. *Clin Infect Dis*. 71:889–890.
- Haehner A, Draf J, Dräger S, de With K, Hummel T. 2020. Predictive value of sudden olfactory loss in the diagnosis of COVID-19. *ORL J Otorhinolaryngol Relat Spec*. 82(4):175–180.
- Hirk R, Hornik K, Vana L. 2020. mvord: an R package for fitting multivariate ordinal regression models. *J Stat Softw*. 93:1–41.
- Hornuss D, Lange B, Schröter N, Rieg S, Kern WV, Wagner D. 2020. Anosmia in COVID-19 patients. *Clin Microbiol Infect*. 26(10):1426–1427.
- Huang C, Huang L, Wang Y, Li X, Ren L, Gu X, Kang L, Guo L, Liu M, Zhou X, et al. 2021. 6-month consequences of COVID-19 in patients discharged from hospital: a cohort study. *Lancet*. 397(10270):220–232.
- Hummel T, Rissom K, Reden J, Hähner A, Weidenbecher M, Hüttenbrink KB. 2009. Effects of olfactory training in patients with olfactory loss. *Laryngoscope*. 119(3):496–499.
- Hummel T, Sekinger B, Wolf SR, Pauli E, Kobal G. 1997. 'Sniffin' sticks': olfactory performance assessed by the combined testing of odor identification, odor discrimination and olfactory threshold. *Chem Senses*. 22(1):39–52.
- Keller A, Malaspina, D. 2013. Hidden consequences of olfactory dysfunction: a patient report series. *BMC Ear Nose Throat Disord*. 13:1–20.

- Klopfenstein T, Kadiane-Oussou NJ, Toko L, Royer PY, Lepiller Q, Gendrin V, Zayet S. 2020. Features of anosmia in COVID-19. *Med Mal Infect.* 50(5):436–439.
- Kohli P, Soler ZM, Nguyen SA, Muus JS, Schlosser RJ. 2016. The association between olfaction and depression: a systematic review. *Chem Senses.* 41(6):479–486.
- Landis BN, Hummel T, Hugentobler M, Giger R, Lacroix JS. 2003. Ratings of overall olfactory function. *Chem Senses.* 28(8):691–694.
- Landis BN, Stow NW, Lacroix JS, Hugentobler M, Hummel T. 2009. Olfactory disorders: the patients' view. *Rhinology.* 47(4):454–459.
- Lechien JR, Chiesa-Estomba CM, Beckers E, Mustin V, Ducarme M, Journe F, Marchan A, Jouffe L, Barillari MR, Cammaroto G, Circiu MP, Hans S, Saussez S. 2021. Prevalence and 6-month recovery of olfactory dysfunction: a multicentre study of 1363 COVID-19 patients. *J Intern Med.* doi: 10.1111/joim.13209.
- Lechien JR, Chiesa-Estomba CM, De Siaty DR, Horoi M, Le Bon SD, Rodriguez A, Dequanter D, Blecic S, El Afia F, Distinguin L, et al. 2020. Olfactory and gustatory dysfunctions as a clinical presentation of mild-to-moderate forms of the coronavirus disease (COVID-19): a multicenter European study. *Eur Arch Otorhinolaryngol.* 277(8):2251–2261.
- Lechien JR, Chiesa-Estomba CM, Hans S, Saussez S. 2021. Pathophysiological mechanisms and management of patients with long-time anosmia related to COVID-19. *J Intern Med.* doi: 10.1111/joim.13257.
- Leopold D. 2002. Distortion of olfactory perception: diagnosis and treatment. *Chem Senses.* 27(7):611–615.
- Levinson R, Elbaz M, Ben-Ami R, Shasha D, Levinson T, Choshen G, Petrov K, Gadoth A, Paran Y. 2020. Time course of anosmia and dysgeusia in patients with mild SARS-CoV-2 infection. *Infect Dis (Lond).* 52(8):600–602.
- Li J, Long X, Zhang Q, Fang X, Li N, Fedorova B, Hu S, Li J, Xiong N, Lin Z. 2021. Tobacco smoking confers risk for severe COVID-19 unexplainable by pulmonary imaging. *J Intern Med.* 289(4):574–583.
- Manesse C, Bellil D, Ferdenzi C, Rouby C, Faure F, Bensafi M. 2018. *A new training protocol to improve olfaction and quality of life in patients with dysosmia: a proof of concept study.* Paris (France).
- Manesse C, Ferdenzi C, Mantel M, Sabri M, Bessy M, Fournel A, Faure F, Bellil D, Landis BM, Hugentobler M, et al. 2021. The prevalence of olfactory deficits and their effects on eating behavior from childhood to old age: a large-scale study in the French population. *Food Qual Prefer.* 93:104273.
- Manesse C, Ferdenzi C, Sabri M, Bessy M, Rouby C, Faure F, Bellil D, Jomain S, Landis B, Hugentobler M, et al. 2017. Dysosmia-associated changes in eating behavior. *Chemosens Percept.* 10:104–113.
- Martin GN, Apena F, Chaudry Z, Mulligan Z, Nixon C. 2001. The development of an attitudes towards the Sense of Smell Questionnaire (SoSQ) and a comparison of different professions' responses. *North Am J Psychol.* 3:491.
- Meinhardt J, Radke J, Dittmayer C, Franz J, Thomas C, Mothes R, Laue M, Schneider J, Brünink S, Greuel S, et al. 2021. Olfactory transmucosal SARS-CoV-2 invasion as a port of central nervous system entry in individuals with COVID-19. *Nat Neurosci.* 24(2):168–175.
- Menni C, Valdes AM, Freidin MB, Sudre CH, Nguyen LH, Drew DA, Ganesh S, Varsavsky T, Cardoso MJ, El-Sayed Moustafa JS, et al. 2020. Real-time tracking of self-reported symptoms to predict potential COVID-19. *Nat Med.* 26(7):1037–1040.
- Moein ST, Hashemian SM, Mansourafshar B, Khorram-Tousi A, Tabarsi P, Doty RL. 2020. Smell dysfunction: a biomarker for COVID-19. *Int Forum Allergy Rhinol.* 10(8):944–950.
- Nordin S, Bende M, Millqvist, E. 2004. Normative data for the chemical sensitivity scale. *J Environ Psychol.* 24:399–403.
- Parker JK, Kelly CE, Gane SB. 2021. Molecular mechanism of parosmia. *MedRxiv*, preprint: not peer reviewed. doi: 10.1101/2021.02.05.21251085.
- Parma V, Ohla K, Veldhuizen MG, Niv MY, Kelly CE, Bakke AJ, Cooper KW, Bouysset C, Pirastu N, Dibattista M, et al. 2020. More than smell – COVID-19 is associated with severe impairment of smell, taste, and chemesthesis. *Chem Senses.* 45:609–622.
- Pfefferbaum B, North CS. 2020. Mental health and the COVID-19 pandemic. *N Engl J Med.* 383:510–512.
- Philpott CM, Boak D. 2014. The impact of olfactory disorders in the United Kingdom. *Chem Senses.* 39(8):711–718.
- R Core Team. 2021. *R: a language and environment for statistical computing.* Vienna (Austria): R Foundation for Statistical Computing.
- Rousseeuw PJ, Hubert M. 2011. Robust statistics for outlier detection. *WIREs Data Mining Knowl Discov.* 1:73–79.
- Satarker S, Nampoothiri M. 2020. Involvement of the nervous system in COVID-19: the bell should toll in the brain. *Life Sci.* 262:118568.
- Schielzeth, H. 2010. Simple means to improve the interpretability of regression coefficients. *Methods Ecol Evol.* 1:103–113.
- Speth MM, Singer-Cornelius T, Oberle M, Gengler I, Brockmeier SJ, Sedaghat AR. 2020. Olfactory dysfunction and sinonasal symptomatology in COVID-19: prevalence, severity, timing, and associated characteristics. *Otolaryngol Head Neck Surg.* 163(1):114–120.
- Vaira LA, Salzano G, Deiana G, De Riu G. 2020. Anosmia and ageusia: common findings in COVID-19 patients. *Laryngoscope.* 130(7):1787.
- Vennemann MM, Hummel T, Berger K. 2008. The association between smoking and smell and taste impairment in the general population. *J Neurol.* 255(8):1121–1126.
- Watabe-Rudolph M, Begus-Nahrman Y, Lechel A, Rolyan H, Scheithauer MO, Rettinger G, Thal DR, Rudolph KL. 2011. Telomere shortening impairs regeneration of the olfactory epithelium in response to injury but not under homeostatic conditions. *PLoS One.* 6(11):e27801.
- Willi S, Lüthold R, Hunt A, Hänggi NV, Sejdiu D, Scaff C, Bender N, Staub K, Schlagenhauf P. 2021. COVID-19 sequelae in adults aged less than 50 years: a systematic review. *Travel Med Infect Dis.* 40:101995.
- Yan CH, Faraji F, Prajapati DP, Ostrander BT, DeConde AS. 2020. Self-reported olfactory loss associates with outpatient clinical course in COVID-19. *Int Forum Allergy Rhinol.* 10(7):821–831.
- Yee TW. 2021. *VGAM: vector generalized linear and additive models. R package version 1.1–5.* Available from: <https://CRAN.R-project.org/package=VGAM>.
- Zeileis A, Hothorn, T. 2002. Diagnostic checking in regression relationships. *R News.* 2:7–10.

Appendix - Online questionnaire

Section 1. Chemosensory disorders

Participants were asked about their perceptions of smells, tastes and trigeminal sensations using the following questions. We also asked the participants whether they consulted a health practitioner specifically for their chemosensory disorders; Which category of practitioner (GP, specialist; which specialist); What were they told and what exam/test were they given; Were they explicitly diagnosed as having dysosmia or dysgeusia (= smell or taste disorders).

A. Your sense of smell

- (1) *In the last days/weeks, have you had any difficulties perceiving odors (e.g., food, coffee, perfumes, or soap)? No/Yes*
- (2) *Is this loss of smell: Mild, Partial / Total?*
- (3) *Was the onset: Sudden/Progressive?*
- (4) *When did this change occur? [date]*
- (5) *Is this condition permanent or do you occasionally experience normal odor perception (fluctuating)? Permanent/Fluctuating*
- (6) *Do you know the cause of this change? Yes/No. If yes, specify the origin: Sino-nasal origin (nasal obstruction due to rhinitis, sinusitis, nasal polyp)/Traumatic origin (following a shock, an accident)/Infectious origin (following a viral/bacterial/fungal infection)/Congenital origin (since birth)*
- (7) *Has this change disappeared since? Yes/No. If yes, when did it disappear? [date]*
- (8) *How did you recover? With surgery/With medication (specify)/With training/With other treatments (specify)/Without treatment*
- (9) *In the past few days/weeks, have smells seemed different from what they usually are (i.e., they don't smell the same)? No/Yes*
- (10) *In the past few days/weeks, have you had any olfactory hallucinations (phantom smells)?
No/Yes (describe)*
- (11) *During these odor changes, did you have a blocked nose? No/Yes*

B. Your sense of taste

- (12) *In the last few days/weeks, have you experienced any changes in taste perception (sweet, salty, sour, bitter)? No/Yes (specify: Difficulty perceiving tastes/Tastes seem different than usual/I have phantom tastes (describe)).*

C. Your trigeminal perceptions

(13) In the last few days/weeks, have you had any changes in the perception of pungency/irritation/cold/hot (in the nose and mouth)? No/Yes (specify: Difficulty in perceiving spiciness-irritancy-cold-hotness/The spiciness-irritancy-cold-hotness seems different from usual/I have phantom sensations of prickly/irritating/cold/hot (describe)).

D. About your sensory changes

(14) Have you consulted a physician about your sensory changes? No/Yes. If yes:

- (a) Who did you consult? (several responses allowed) A general practitioner/A specialist (specify which specialty)*
- (b) What were you told, and what exam/test were you given?*
- (c) Have you been explicitly diagnosed as dysosmic or dysgeusic (= with smell or taste disorders)? No/Yes*

Section 2. Quality of life

Participants were first asked whether chemosensory disorders were incapacitating. Then, they were questioned about 16 feelings or habits related to food, social interactions, danger detection and enjoyment of the environment (items adapted from Manesse et al., 2018). Participants answered “more than usual”, “less than usual” or “neither more nor less than usual” to the following questions:

“During your sensory modifications,

- (1) Did you enjoy eating*
- (2) Did you let your meals burn in the oven or on the stove*
- (3) Did you wear perfume*
- (4) Did you add sugar to your meals to make them match up to your taste*
- (5) Did you prefer eating alone*
- (6) Did you have domestic accidents*
- (7) Did you add salt to your meals to make them match up to your taste*
- (8) Did you seek to control your body odors (in order to not bother others, for example)*
- (9) Did you detect the odor of smoke and gas*
- (10) Did you add spicy condiments (chili pepper, mustard) to your meals to make them match up to your taste*
- (11) Did you enjoy smelling flowers' perfume*
- (12) Did you accidentally eat rotten food*
- (13) Did you add fatty condiments (mayonnaise, oil) to your meals to make them match up to your taste*
- (14) Did you take showers*
- (15) Did you burn your clothes while ironing them*
- (16) Did you seek to smell the odors of your close ones (children, partner)”*

Annexe 3 : « Perte de l'odorat comme indicateur précoce de l'évolution de la pandémie de COVID-19 » (publié)

Smell and taste changes are early indicators of the COVID-19 pandemic and political decision effectiveness

Denis Pierron ^{1,13,14}✉, Veronica Pereda-Loth ^{1,13,14}, Marylou Mantel ², Maëlle Moranges ², Emmanuelle Bignon³, Omar Alva¹, Julie Kabous¹, Margit Heiske¹, Jody Pacalon³, Renaud David⁴, Caterina Dinnella⁵, Sara Spinelli⁵, Erminio Monteleone⁵, Michael C. Farruggia ⁶, Keiland W. Cooper ⁷, Elizabeth A. Sell ⁸, Thierry Thomas-Danguin ⁹, Alyssa J. Bakke ¹⁰, Valentina Parma ¹¹, John E. Hayes¹⁰, Thierry Letellier¹, Camille Ferdenzi ^{2,13,14}, Jérôme Golebiowski ^{3,12,13,14}✉ & Moustafa Bensafi ^{2,13,14}✉

In response to the COVID-19 pandemic, many governments have taken drastic measures to avoid an overflow of intensive care units. Accurate metrics of disease spread are critical for the reopening strategies. Here, we show that self-reports of smell/taste changes are more closely associated with hospital overload and are earlier markers of the spread of infection of SARS-CoV-2 than current governmental indicators. We also report a decrease in self-reports of new onset smell/taste changes as early as 5 days after lockdown enforcement. Cross-country comparisons demonstrate that countries that adopted the most stringent lockdown measures had faster declines in new reports of smell/taste changes following lockdown than a country that adopted less stringent lockdown measures. We propose that an increase in the incidence of sudden smell and taste change in the general population may be used as an indicator of COVID-19 spread in the population.

¹Équipe de Médecine Evolutive Faculté de chirurgie dentaire; UMR5288; CNRS/Université Paul-Sabiater Toulouse III, Toulouse 31400, France. ²Lyon Neuroscience Research Center, CNRS UMR5292, INSERM U1028, University Claude Bernard Lyon 1, Bron, France. ³Université Côte d'Azur, CNRS, Institut de Chimie de Nice UMR7272, Nice, France. ⁴Université Côte d'Azur, CHU de Nice, Nice Memory Clinic, Nice, France. ⁵University of Florence, Florence, Italy. ⁶Interdepartmental Neuroscience Program, Yale University, 333 Cedar Street, New Haven, CT 06520, USA. ⁷Department of Neurobiology and Behavior, University of California, Irvine, CA 92697, USA. ⁸Perelman School of Medicine, University of Pennsylvania, 3400 Civic Center Blvd, Philadelphia, PA 19104, USA. ⁹Centre des Sciences du Goût et de l'Alimentation, INRAE, CNRS, AgroSup-Dijon, University Bourgogne Franche-Comté, Dijon, France. ¹⁰The Pennsylvania State University, Philadelphia, PA 19104, USA. ¹¹Temple University, Philadelphia, PA 19122, USA. ¹²Department of Brain and Cognitive Sciences, Daegu Gyeongbuk Institute of Science and Technology, Daegu 711-873, South Korea. ¹³These authors contributed equally: Denis Pierron, Veronica Pereda-Loth, Camille Ferdenzi, Jérôme Golebiowski, Moustafa Bensafi. ¹⁴These authors jointly supervised this work: Denis Pierron, Veronica Pereda-Loth, Camille Ferdenzi, Jérôme Golebiowski, Moustafa Bensafi. ✉email: denis.pierron@univ-tlse3.fr; jerome.golebiowski@univ-cotedazur.fr; moustafa.bensafi@cnrs.fr

ollowing similar decisions in China and Italy, a strict lockdown was enforced in France beginning on March 17, 2020 to block the progression of COVID-19 and alleviate pressure on hospitals. One issue currently faced by governments is how to conduct the progressive relaxation of the lockdown¹, which needs to be conducted systematically and carefully to prevent subsequent outbreaks while facilitating economic activity and recovery. On May 7, 2020, the French government categorized each geographical area as being red or green, depending on their COVID-19 prevalence. Compared to green areas, red areas were characterized by: (i) higher active circulation of the virus, (ii) higher level of pressure on hospitals (i.e., CCRU occupancy), and (iii) reduced capacity to test new cases (Fig. 1a). In each area, red/green labels were used to define steps associated with the local relaxation of lockdown. The French Ministry of Health used the ratio of consultations for suspected cases of COVID-19 to general consultations at the emergency room (ER) in hospitals as an indicator to assess the active circulation of the virus (detailed in “Methods” section). Concurrently, changes in smell and taste are prominent symptoms of COVID-19^{2–5}, as has consistently been demonstrated in many countries (e.g., Iran⁶, Spain⁷, France⁸, Italy⁹, Germany¹⁰, and the UK², among others). More critically, these chemosensory changes generally occur earlier than other symptoms⁹ and may constitute more specific symptoms than fever or dry cough^{2,11}. Accordingly, monitoring self-reported changes in smell and taste could thus provide early and specific information on the spread of COVID-19 in the general population and support health system monitoring to avoid daily CCRU admission overflows. Using data from a global, crowd-sourced study deployed in 30+ languages (Global Consortium for Chemosensory Research survey, GCCR, see “Methods” section), we tested whether changes in smell/taste at the population level could be used as an early indicator for local COVID-19 outbreaks. As pre-registered (see “Methods” section), our primary aim was to test the association between self-reported smell and taste changes and indicators of pressure in hospitals (COVID-related hospitalizations, CCRU admissions, and mortality rates) for each French administrative region over the last 3 months. Our secondary aim was to examine temporal relationships between the peak of smell and taste changes in the population and the peak of COVID-19 cases and the application of lockdown measures. The potential for self-reported smell and taste loss to serve as an early indicator of the number of COVID-19 cases—and hence hospital stress—was tested in a natural experiment by comparing France with Italy and the UK, which implemented lockdown with different timing and levels of stringency. Here, we show that self-reports of smell/taste changes are closely associated with hospital overload and are early markers of the spread of infection of SARS-CoV-2.

Results

Changes in smell and taste are associated with overwhelmed healthcare systems. The relationship between self-reported changes in smell and taste by French residents (diagnosed as COVID-19+ or not, see “Methods” section and Supplementary Table 1) and estimators of local healthcare system stress was evaluated geographically. Figure 1a depicts the geographical distribution in red and green regions (as defined by the French government) and participants who self-reported changes in their smell and taste. Red areas of France account for 40.8% of the population. Green areas are clustered into a group with both a low number of self-reported chemosensory changes and a low number of admissions to CCRUs (Fig. 1b). Red areas show an opposite trend (Chi-square $<1 \times 10^{-200}$ and Biserial correlations $p < 1.3 \times 10^{-2}$). A strong relationship exists between self-reported changes in smell and taste and the number of admissions to

CCRUs ($R_{\text{smell}} = 0.88$, $p = 8.9 \times 10^{-08}$). This correlation remained significant even after removing the two most impacted areas (Alsace and Ile de France, $R_{\text{smell}} = 0.72$; $p < 3 \times 10^{-04}$), indicating that the significant relationship is not driven solely by these two regions.

Strikingly, use of self-reported chemosensory changes produced a stronger correlation than the current governmental indicator of virus circulation (Fig. 1c). Overall, smell/taste changes are better correlated with the number of COVID-19 admissions to hospitals than the current governmental indicator i.e., the ratio of ER consultations for suspicion of COVID-19 to general ER consultations ($R_{\text{smell}} = 0.81$, $p = 6.71 \times 10^{-06}$ vs. $R_{\text{gov}} = 0.60$, $p = 3.35 \times 10^{-03}$); the same pattern was found for the number of COVID-19 related deaths ($R_{\text{smell}} = 0.75$, $p = 5.62 \times 10^{-05}$ vs. $R_{\text{gov}} = 0.58$, $p = 4.97 \times 10^{-03}$ see Supplementary Table 2). Further, when smaller geographical areas were considered (France is divided into 96 administrative units, called departments), these correlations remained highly significant (e.g., admissions to CCRUs: $R_{\text{smell}} = 0.76$, $p < 5 \times 10^{-19}$) (Fig. 1c). Moreover, the three relationships (change in smell/taste versus COVID-19-related hospitalization, resuscitations, and death) also remained highly significant when considering only individuals who were not clinically diagnosed by a medical professional but considering themselves showing some symptoms of COVID-19 (e.g., admissions to CCRUs: $R_{\text{smell}} = 0.83$, $p = 1.65 \times 10^{-06}$). Potential sampling bias due to regional media coverage of our survey (Supplementary Table 3) and self-reported chemosensory changes by region was ruled out by confirming these variables were not correlated ($R < 0.01$, $p > 0.9$).

Notably, relationships between pandemic markers and online searches related to chemosensation were also significant in France. Google queries related to smell or taste loss (“perte odorat,” “perte goût” in French) were correlated with the three measures of an overwhelmed healthcare system described above (e.g., CCRU admissions: $R_{\text{smell}} = 0.8$, $p < 4 \times 10^{-03}$, see Supplementary Table 2).

Changes in smell and taste are early markers of the effectiveness of political decisions. Next, we examined the temporal dynamics in France of self-reported changes in smell/taste, the current governmental indicator (ratio of ER consults), and the number of CCRU admissions due to COVID-19 before and after the lockdown period. As shown in Fig. 1d, the peak of the onset of changes in smell/taste appeared 4 days after the lockdown and for these individuals, the first reported COVID-19 symptoms occur even earlier. Conversely, the governmental indicator of ER consults only peaked 11 days after the lockdown, while the peak of CCRU admissions was shifted later by 14 days. This is consistent with emerging data showing that COVID-19-related changes in smell and taste occur in the first few days after infection^{6,12–14}. The robustness of smell and taste changes over time was assessed in two ways. First, we showed the peak of smell/taste changes remained the same regardless of our survey’s completion date (Supplementary Fig. 1a). Second, we observed the exact same peak when analyzing a separate French survey performed on 950 individuals and focusing on smell alterations in the French population independently of COVID-19 (see “Methods” section): the peak of olfactory changes again occurred 4 days after the lockdown decision, and this was independent of survey completion dates (Supplementary Fig. 1B). The robustness of smell and taste changes was also observed over age (Supplementary Fig. 2A) and gender (Supplementary Fig. 2B). Finally, we also show that the observed peak does not correspond to seasonal occurrence of allergies in France based on the ratio of consultations for Allergy to general consultations at the emergency room (Supplementary Fig. 3).

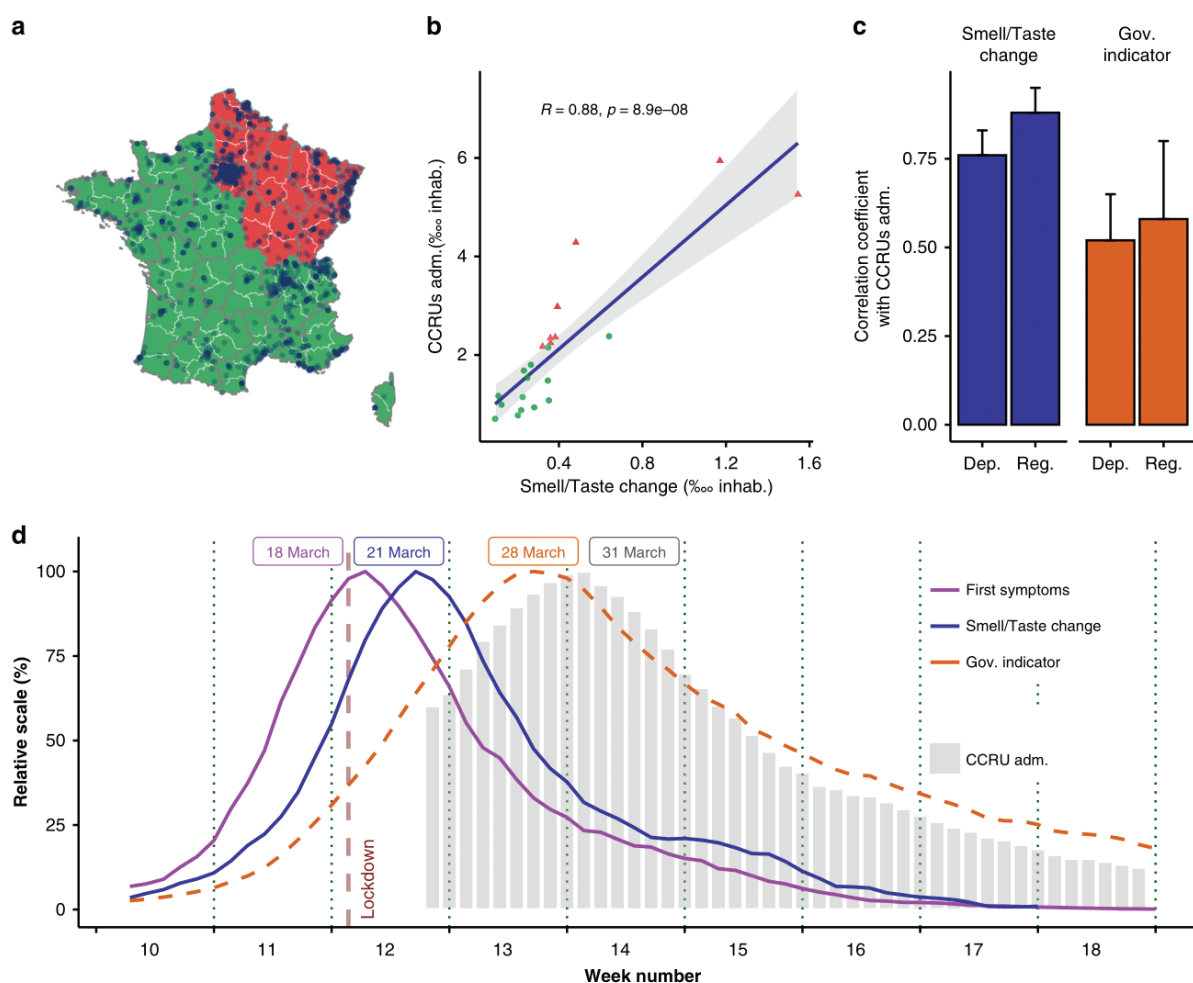


Fig. 1 Changes in smell and taste as indicators of overwhelmed healthcare systems: geographic and time-related approaches. **a** French regions were assigned a green or red status by the French government to guide local relaxation of lockdown protocols. Dots represent people self-reporting smell and taste changes in a web-based survey. Base map is from OpenStreetMap and OpenStreetMap Foundation. **b** The number of COVID-19-related CCRU admissions (as of May 11, 2020) correlated with the number of self-reported chemosensory changes (between March 1 and May 11, 2020, total $n = 3832$). Green dots correspond to regions with a post-lockdown level labeled green, and red triangles indicate regions considered red. Values are standardized based on the number of inhabitants (inhab.) for each regions. The two red triangles with CCRU admissions >5 are Alsace and Ile de France. The gray band represent the confidence interval of the linear smooth (formula 'y - x') R and p represent value of the test for association between paired samples, using one of Pearson's product moment correlation coefficient, without correction for multiple comparisons. **c** Colored bar represent the value of computed correlation coefficients (confidence intervals are depicted as thin black bars) between the number of CCRU admissions per area and i) the number of people reporting smell and taste changes ($n = 3832$, blue), and ii) the governmental indicator (Gov. indicator), ratio of ER consults for COVID-19 (orange). Analyses were done both at the level of metropolitan regions (Reg) and departments (Dep). **d** Temporal relationships in France between smell/taste change symptom onset (blue solid line, $n = 1476$), the governmental indicator (orange dashed line), and COVID-19 cases in CCRUs (gray bars) around the lockdown period (vertical dashed line). Data are 7-day running averages, normalized to the day with the highest value.

Further, analyses of Google searches confirm this temporal relationship: on the same days where survey participants report experiencing their first symptoms (around March 18, 2020), there was a peak of Google queries for terms associated with early COVID-19 symptoms (fever, cough, aches, Supplementary Fig. 4A). A few days later, the peak of online queries for "taste loss" and "smell loss" is synchronized with the report of smell and taste changes (Supplementary Fig. 4B). One week later, queries for shortness of breath preceded the peak of CCRU admissions (Supplementary Fig. 4C). Collectively, these results indicate a significant fraction of French COVID-19 patients followed the same symptom time course, experiencing initial symptoms at the very start of the lockdown, which might be representative of a peak of infection a few days before the lockdown. This is

consistent with the ultimate goal of the lockdown, which was to decrease the number of new infections following implementation. Thus, the period immediately prior to lockdown represents the expected peak of new infections. In France, a large population may have been infected two days before lockdown because that weekend was crowded and sunny and occurred over the course of election day. Further, there were busier train stations and supermarkets in anticipation of a shortage of supplies during lockdown¹².

These data suggest that the short-term efficacy of a lockdown could be monitored by tracking changes in smell and taste in the population. To assess whether such a prediction might generalize to other countries, we performed parallel analyses with data from Italy and the UK, where the lockdown measures were established

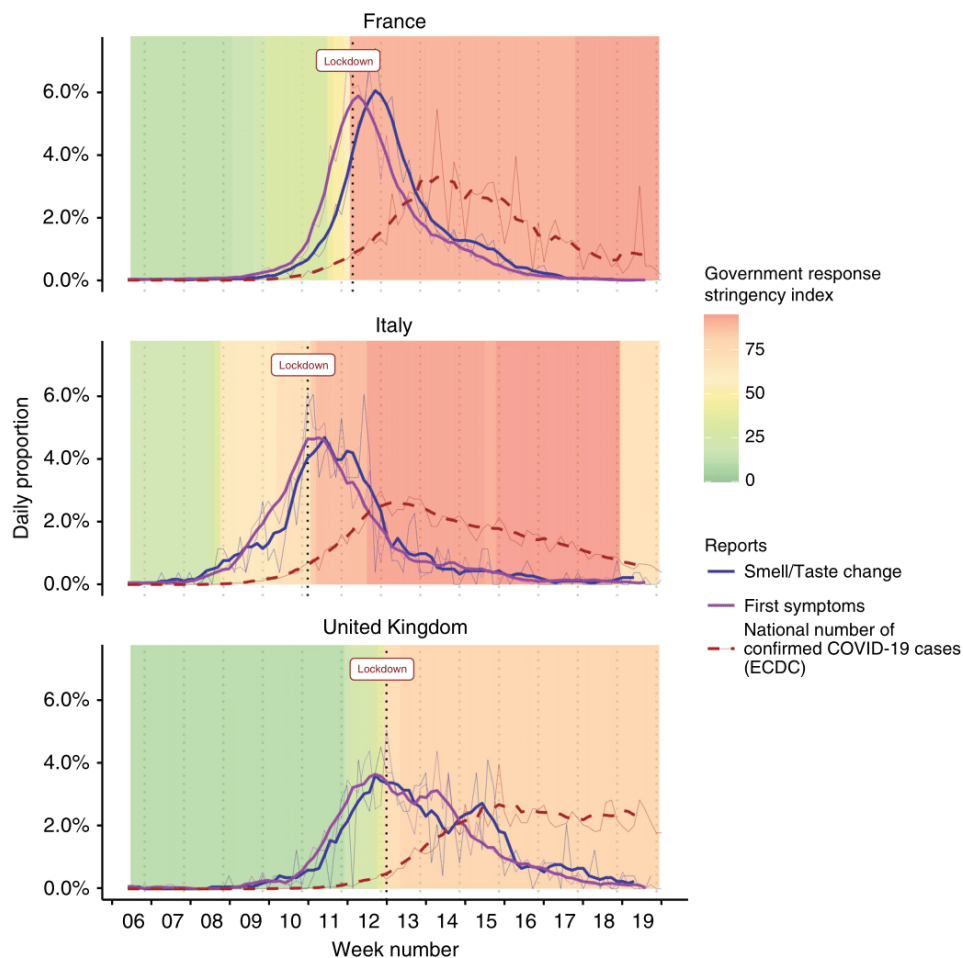


Fig. 2 Evolution of COVID-19 indicators before and after the lockdown in France, Italy and the UK. The daily proportion of first symptoms is shown as a violet line (France, $n = 4720$, Italy, $n = 1241$, UK, $n = 750$). The daily proportion of smell/taste changes is shown as a blue line (France, $n = 1487$, Italy, $n = 264$, UK, $n = 263$). The daily proportion of COVID-19 confirmed cases from the European Centre for Disease Prevention and Control (ECDC) is shown as a red dashed line. Each panel shows both raw data (thin line) and the corresponding 7-day running average (thick line). The government response stringency index is shown as the background color.

with different levels of severity (see Fig. 2). We monitored the dynamics of confirmed COVID-19 cases, self-reported first symptoms, and self-reported taste and smell changes, and compared them as a function of the governmental stringency index. Immediately after lockdown, we found that the two countries with the higher stringency index experienced a more rapid decrease in both self-reported smell and taste changes and COVID-19 symptoms. Further, as expected, the evolution of confirmed COVID-19 cases differs according to the stringency index. The governments of Italy and France rapidly increased their stringency index, which led to a sharp decrease in COVID-19 symptoms and cases. In contrast, in the UK, the number of people in the UK reporting symptoms showed a slower decrease, presumably due to a less severe lockdown policy, and the number of confirmed cases remained high during the observation window. In each country, self-reported smell and taste changes can be regarded as a useful metric to predict the dynamics of confirmed COVID-19 cases. That is, when the number of new onsets of chemosensory changes decreases sharply (France and Italy), the number of confirmed COVID-19 cases also decreases, albeit with a lag of two weeks. On the contrary, a slow decrease in the number of new onset chemosensory changes is associated with a plateau of confirmed cases (UK).

Discussion

The present analyses reveal a strong spatial and temporal relationships between self-reported smell and taste changes and multiple indices of health care system stress, such as admissions to CCRUs. This is consistent with cumulative evidence showing a high prevalence of chemosensory alterations in patients affected by COVID-19 in Europe (France^{8,14}, Italy⁹, UK^{2,15,16}). Participants endorsed smell and taste changes only 3–4 days after their first symptoms. Such early chemosensory estimators may represent a cost-effective and easy way to implement alternative surveillance methods to large-scale virology tests, which are difficult to perform, costly, and time-consuming, especially during a pandemic.

A prominent question raised by these findings is whether the smell and taste changes observed in our study are solely related to COVID-19 or whether they can be explained by other temporal patterns, like seasonal illnesses or allergies. To the best of our knowledge, there are no existing studies that have explored the dynamics of sudden anosmia (as in COVID-19) throughout the year in France. Relationship between olfactory disturbances and seasons have been reported in Korea, Germany or US with a moderate increase of anosmia prevalence in spring^{17–19}. Although the cyclical pattern of smell/taste changes might

overlap, the amplitude of reported changes (either due to allergy or viral affection) were very limited compared to the present report. To further rule out the possibility, we examined whether the annual peak of allergies in France could explain the peak of smell and taste changes observed here. In analyzing existing French governmental data, we found that the annual peak of allergies in France occurred around week 30 (beginning of summer), multiple weeks after the observation window of the present study (from week 5 to week 20, Supplementary Fig. 3). Further, the French national aerobiological surveillance network (RNSA, <https://pollens.fr>), which follows pollen concentration in the atmosphere, has also indicated the first week of lockdown was very low risk for seasonal allergies. In addition, when considering Google Trends data, we did not observe any similar peaks in queries for smell/taste loss in the corresponding time period in previous years. Finally, a comparative study in Israel²⁰ showed that in COVID-19 suspected patient the frequency of smell change is almost ten times higher in a COVID-19 positive patients (68%) than in COVID-19 negative (8%). Considering that most of the participants of the present study are diagnosed with COVID-19 and that their description of a sudden loss of smell/taste is consistent with the now typical presentation of COVID-19 symptoms, it is highly probable that COVID-19 infection is the main reason of their smell and taste change. Collectively, these data suggest the peak of smell and taste changes studied here are more consistent with sudden COVID-19 viral infections rather than an artifact due to seasonal illnesses.

The time lag between the onset of COVID-19-related symptoms and their declaration by the respondents of our study also deserves comment. Although immediate reporting of symptoms would have been ideal, such reporting is not possible within the context of the sudden first wave of a new viral pandemic. A similar time lag has been observed in other large-scale studies focusing on olfaction and COVID-19²¹. Indeed, this time lag is inevitable given the preparation time required for scientists and clinicians design and launch such a survey, with appropriate ethics approval, once anosmia and ageusia began to emerge as cardinal symptoms of COVID-19. The vast majority of participants completed the survey between April 10th and April 19th, 2020, and most of them declared a date of onset of their symptoms roughly a month earlier (although a small fraction of participants did indicate onset prior to 2020). A possible consequence of a time lag between survey completion and the effective date of symptom onset is that subjects' statements may have been influenced by major societal events such as the lockdown decision, potentially creating some recall bias. To examine whether the date of a major event like the lockdown might bias dates of reported smell and taste loss, we explored narrative descriptions provided by our participants. By analyzing responses to the optional open-ended question "Please describe the progression or order you noticed your symptoms", we observed that, for France, a mere 11 of 3705 people (who have filled the optional question) used the term "confinement" ("lockdown") in their description of the onset date. Separately, another factor that mitigates concerns about a potential recall bias is the stable nature of participant's statements, regardless of their date of completion. That is, logic suggests, the longer the time between the onset date of smell and taste loss and the reporting date, the greater the recall bias should be. However, our data clearly show that regardless of the date of completion, the onset date falls within the same period (Supplementary Fig. 1). Finally, other evidence against a potential recall bias comes from Google Trends data. Analyzing real-time Google queries in March, we observed a very particular trend in France (Supplementary Fig. 4). We first observed a peak of queries for terms associated with early COVID-19 symptoms (fever, cough, aches) synchronized with

the declared onset of the first symptoms in the survey (around March 18th). A few days later, a peak of online queries for "taste loss" and "smell loss" was seen, and this was synchronized with the date reported of smell and taste changes in our survey. The striking concurrence between Google queries and reports in our survey argues against the idea that a recall bias could be driving the effects described here.

Another important factor to consider in our survey is the way the press and media might have influenced our findings. Indeed, when the survey was launched, smell and taste changes were reported as symptoms of COVID-19 in the national and local media, which might have influenced respondents to remind themselves of such symptoms and to then report these changes on the survey. Such an emphasis on smell and taste loss would have biased attempts to explore the prevalence of chemosensory deficits in COVID-19. However, the primary aim of the present investigation was not to focus on the prevalence of anosmia and ageusia with COVID-19, but rather to explore use of reported smell and taste loss as indicators of COVID-19 pandemic. Still, the media coverage of our survey could also have biased the selection of participants geographically, as some French regions received more media coverage than others. However, as reported above, there was no correlation between the number of participants in a given region and the intensity of media and press coverage for the survey in that same region. Finally, when participants were asked to describe the chronology of their symptoms, they did not refer to the media coverage as a prominent element influencing their awareness of their smell/taste changes. While this does not exclude an implicit and non-verbalized bias due to media coverage, this pattern suggests a genuine report of symptoms with a high occurrence of COVID symptoms just after the lockdown.

An interesting question raised by our findings is what impact they might have on government strategies in a pandemic. Following lockdown, the rapid decrease of self-reported changes in smell and taste in France may be representative of the effectiveness of this decision in reducing infection rates. Similarly, data from Italian participants show highly similar patterns, but with a one-week difference compared to the French data. This might reflect highly similar responses by the Italian and French governments. Conversely, the prevalence of chemosensory changes in the UK shows a more gradual decrease. The UK government began with advice to avoid pubs, clubs and theaters, and to work from home from March 16, with restrictions around March 18. However, a lockdown was not declared until March 23, and this was less stringent than those in France or Italy. Notably, new COVID-19 cases in the UK showed a plateau phase which is not observed in either France or Italy. Accordingly, we conclude that collecting online information about changes in smell and taste from residents (even retrospectively) may be a valuable metric of the effectiveness of reopening strategies related to the COVID-19 pandemic.

Practically, in areas where smell and taste changes are notable COVID-19 symptoms, the proportion of individuals who self-report changes in their ability to smell or taste might be an early indicator of subsequent demand for healthcare. If confirmed, continuous monitoring of changes in smell and taste perception would then be a highly cost-effective, minimally invasive, and reliable way to track future COVID-19 outbreaks. When used this way, we caution that particular attention must be paid to potential selection bias. That is, self-report studies online can be impacted by multiple selection biases, including socioeconomic status, fluency with technology and willingness and interest in participating in scientific research. When considering the present data, at least 3 parameters may contribute to a selection bias in our sample: (1) the age, (2) the gender of the participants, and (3)

the format and the advertising of the survey. Regarding participant age, our study cohort (mean 40.7 years, $sd = 12.4$) was quite similar to the French population mean (41.1 years, according to INSEE, <https://www.insee.fr/fr/statistiques/1893198>); however, we did only include individuals over 18 due to issues of consent, and administrative reasons, and seniors were also less represented. For gender, our sample contained a greater proportion of women (67%) compared to men, which might influence the results. However, additional analysis showed no differences in peaks of smell/taste changes across age or gender, minimizing concerns that such selection biases may have influenced present results (See Supplementary Fig. 2). We also tested the potential selection bias due to format and the advertising of the survey, by comparing the GCCR dataset with an independent second study performed on French residents (see “Methods” section). Remarkably we observed highly similar results across studies where advertising, inclusion criteria, and survey format were different.

Based on the present findings, we highlight the paramount importance and robustness of associations between smell/taste changes and COVID-19 and we strongly endorse the need for additional large-scale validation studies to assess the causality between the observed association between smell/taste changes and indicators of the COVID-19 pandemic. This could be achieved by setting up a simplified interface where selection biases are controlled for (age, gender, motivation, media coverage, socioeconomic level, etc.) through both traditional and online media—and whereby real time information about changes in smell and taste in the general population may be available to decision-makers. Subjects’ participation in the questionnaire and the reliability of the answers should also be considered. In particular, if a participant knows how their answers may influence enforcement of lockdown, their answers might become less truthful. This motivation can be expressed through different forms of behavior. Whereas some individuals may tend to provide statements that minimize their symptoms in order to avoid strict containment measures, others will maximize their declaration to maintain the lockdown, or will provide honest answers in order to participate in the collective effort to better understand the COVID-19 pandemic. These motivational factors are a recurrent risk in online studies and different strategies should be held to control for them in future predictive studies. Based on the above, a large implementation of the study of smell and taste changes in institutional models should allow for monitoring of COVID-19 spread. This might be especially relevant in areas in which testing proves difficult or delayed and for future outbreaks that may overlap with other seasonal viral diseases which share many of the symptoms (fever, cough etc.) but whose treatment or prevention (vaccination) are less demanding in terms of critical care than COVID-19. We advocate that self-report surveys should be used to enhance other strategies such as large-scale PCR tests and COVID-19 symptom assessments (including anosmia and ageusia) in primary/secondary care.

In summary, we propose that an increase in the incidence of sudden smell and taste change in the general population may be used as a valuable minimally invasive indicator of coronavirus spread in the population. To formally test the temporal relationship between chemosensory changes and spread of the disease, we recommend that a large-scale causal study in different countries be conducted on real-time monitoring of self-reported changes in the ability to smell or taste. Such a prospective study will allow for the creation of statistical models that can assist in prediction of future hospital admissions for COVID-19. Further, it could also help decision-makers take important measures at the local level, either in catching new outbreaks sooner, or in guiding the relaxation of local lockdowns, given the strong impact of lockdown on economic and social activities.

Methods

Online survey. This study is mainly based on data from the Global Consortium for Chemosensory Research survey (GCCR, <https://gcchemosensr.org/>) – a global, crowd-sourced online study deployed in 30+ languages²². The data analyzed here were collected from April 7 to May 14, 2020. The protocol complies with the revised Declaration of Helsinki and was approved as an exempt study by the Office for Research Protections at The Pennsylvania State University (Penn State) in the U.S.A. (STUDY00014904; PI Hayes).

Participants in the GCCR questionnaire were recruited by word of mouth, as well as through social and traditional media (flyers, social media, television, radio) during the COVID-19 pandemic. It was well covered by the French press, as over 70 articles mentioned the project, at both the regional and national level (see Supplementary Table 3). Respondents received no monetary incentive for their participation. Inclusion criteria were as follows. (i) Questionnaire completion was allowed only to participants who indicated they had suffered from a respiratory disease in the past two weeks, whether they noticed a change in their taste/smell or not. (ii) Participants aged 18 years old or younger were excluded.

For the analyses conducted in this article, only individuals reporting a change in smell and/or taste perception were included, based on the question “Have you had any of the following symptoms with your recent respiratory illness or diagnosis?”. Moreover, to exclude unreliable entries, participants must have reported a quantitative decrease of at least 5 on a 0-to-100 rating scale between their ability to smell and/or taste before and during their recent respiratory illness or diagnosis. Therefore, Due to this inclusion criteria, “smell/taste change” is equivalent to a quantitative decrease of participant ability to smell and/or taste. We then extracted individuals from the full dataset who reported living in France, Italy or the UK. As the country of residence was completed as a text entry, we allowed for typical variations (e.g., “United Kingdom” or “UK”), spelling mistakes, use of different languages (e.g., “Italie” or “Italia”), as well as subdivisions (e.g., “Scotland”) and major cities (“Paris”). Metropolitan France was split into 13 so-called “regions” in 2016. However, we considered the former system where France was split into 22 regions here, since the organization of the health system mostly remains based on the structure built before 2016. An alternative, finer granularity, splits metropolitan France into 96 so-called “departments.” To retrieve the French department and region of the participants, we used the city of residence they reported in the questionnaire and combined them with the French public website (data.gouv.fr, after a semi-manual correction of spelling). Participants came from all metropolitan departments but three (Mayenne, Creuse, Cantal). Consequently, the number of responses analyzed in France was between $n = 1476$ and 4720 depending on the analysis conducted (i.e., on whether the information of interest was present or missing and the date range of analysis, see Supplementary Table 1 for details). For comparison, between 264 to 1241 participants from Italy and between 243 to 750 participants from the UK were included. Most participants were women (FR:66.38%, IT:69.3%, UK:76.0%), and the mean age was around 40 [FR = 40.7 ($sd = 12.4$), IT = 41.1 ($sd = 11.4$), UK = 41.09 ($sd = 12.1$)]. In the French data, a total of 15% of individuals tested positive for COVID-19 (lab result) and 44% were diagnosed clinically by a medical professional from their symptoms. The remaining 41% were not diagnosed for COVID-19 but declared a change in perception of either smell or taste. The number of participants was normalized by region, by using the number of inhabitants in each region as estimated by the French public statistics office, INSEE. Finally, the time of onset of smell and taste change was assessed via responses to several optional open-ended questions. These included: “Please describe the progression or order you noticed your symptoms” and the time of onset of recent disease by the question: “At what date did you first notice symptoms of your recent respiratory illness? Provide your best guess or leave blank if you do not remember.”

Complementary and independent French Survey. The data of another online survey were used to evaluate the robustness of the temporal evolution of smell and taste changes. This survey was conducted in the French population between April 8 and May 8, 2020 and aimed at characterizing chemosensory disorders in people with and without COVID-19, as well as their consequences on quality of life. The data of 950 respondents were eligible for comparison with data from the GCCR survey, i.e., responses where both the date of completion and the date of smell loss onset were provided. Only responses that were complete and from people who were responding to the questionnaire for the first time and were over age 18 were included. This survey was approved by the CNRS ethics committee. Data collection was strictly anonymous. The protocol complies with the revised Declaration of Helsinki and the study was approved by the ethics committee of the Institute of Biological Sciences of the CNRS on the 3rd of April 2020 (DPO #TRRECH-467). All individuals provided informed consent when participating in the survey.

Online trends. Trends of online queries by French region were performed using Google Trends, a tool returning the popularity of a search term in a specific state or region. Google is by far the most used search engine in France (>90% of internet searches, according to StatCounter Global Stats). We looked for the popularity of terms (listed in Supplementary Fig. 3, using default selection of “All categories” and “Web search”), within the timeframe of February 1, 2020 to May 10, 2020 (from the month of the first official COVID-related death in Europe to the end of lockdown in France). It should be noted that Google Trends does not provide the

actual numbers of searches but rather a relative score from 0 to 100 (100 corresponding to the day with the greatest number of searches during the specified time period). To compare Google Trends scores between French regions, we transformed them by computing the relative number of queries per day in the region of interest. For example, despite a value of 100, the peak day might represent only 5% of the total number of queries related to the topic across the timeframe of interest (see above).

Healthcare system data. The French governmental indicator to estimate the circulation of the virus was calculated from the ratio of consultations for suspected COVID-19 to general consultations at the emergency room (ER) in hospitals. This ratio corresponds to the medical diagnostic for COVID-19 suspicion (codes CIM10: U07.1, U07.10, U07.11, U07.12, U07.14, U07.15, U04.9, B34.2, B97.2). The definition of COVID-19 has evolved rapidly during the lockdown period but the diagnosis is principally based on symptoms of COVID-19 considered as common such as fever, cough, and dyspnea (difficulty breathing). To the best of our knowledge, anosmia and ageusia were officially considered in France as putative symptoms of COVID-19 from a letter of the Direction Générale de la Santé (April 1st) and communication of the Haut Conseil de la Santé Publique (a letter dated April 8, published online April 15, following a letter from the CNP-ORL dated March 20). Areas with values of the French governmental indicator higher than 10% are considered having a high virus circulation. This indicator contributes to the assignment of a red/green label. Allergies incidence in previous years were calculated from the ratio of consultations for Allergy to general consultations at the emergency room (ER) in hospitals.

Data dealing with the health status across countries (number of COVID-19 cases and deaths for each day) were downloaded on May 22, 2020 from the European Centre for Disease Prevention and Control databank (ECDC, <https://www.ecdc.europa.eu/en>). Data regarding healthcare system stress in France (hospitalizations, CCRU entries and deaths) were also downloaded on May 22 from the French Public Health website (Géodes, Santé Publique France, <https://geodes.santepubliquefrance.fr/#c=home>). Here, we use the term CCRU (Critical Care Resuscitation Unit) to translate the French hospital service of "Réanimation." Raw data were normalized across regions with regard to their number of inhabitants as estimated by INSEE. The temporal evolution of the stringency of government response was retrieved from the Oxford COVID-19 (<https://www.bsg.ox.ac.uk/research/research-projects/coronavirus-government-response-tracker>). Here, the stringency level of a country is computed according to which measures of a list of items (e.g., school closures, cancellation of public events, international travel controls, etc.) are undertaken. For the post-lockdown situation, the color assigned by the French government to each department was downloaded on May 12 from the government website. Only data before May 11 (the initial lift of the lockdown) were included in the analyses.

Statistical analyses. Statistical analyses were pre-registered at the Open Science Framework (OSF). Data were analyzed using R software (4.0) and its standard packages (maps, ggplot, etc.). Data were grouped at the national level (France, Italy, UK). In France they were also grouped at the regional level (according to the division into 22 regions in place prior to the 2016 reform). The rationale behind this is that the healthcare system is still structured following this organization, with University Hospitals in regional main cities serving patients of the surrounding departments. Participants from overseas French territories were not included in the geographical analysis because of too few data ($n < 10$). The relationship between (1) GCCR responses (or online queries), and (2) public health data was determined using parametric (e.g., Pearson correlations) statistics as allowed by the normal distribution of the variable of interest. The association between GCCR participant and red/green post-lockdown status was tested using Chi-square tests and Biserial correlations. Complementary analyses not planned in the pre-registration included: (i) the analysis using the independent French online survey (see section "Complementary and independent French Survey" of the methods), (ii) the correlation between regional media coverage and the number of responses to the online survey per region, (iii) the correlation at the level of department, (iv) the correlation excluding extreme points, and (v) the correlation with the government indicator. Pre-registered statistical analyses not presented here include: (i) Mann-Kendall trend test and Change-point Detection test to detect time series changes, and (ii) part of the Google Trends analysis.

Reporting summary. Further information on research design is available in the Nature Research Reporting Summary linked to this article.

Data availability

The authors declare that the data supporting the findings of this study are available within the paper and its supplementary information files. (Source Data file). Source data are provided with this paper.

Code availability

R scripts are available on the osf server (<https://osf.io/gew7p/>).

Received: 5 June 2020; Accepted: 14 September 2020;
Published online: 14 October 2020

References

- Gilbert, M., Dewatripont, M., Muraille, E., Platteau, J.-P. & Goldman, M. Preparing for a responsible lockdown exit strategy. *Nat. Med.* **26**, 643–644 (2020).
- Menni, C. et al. Real-time tracking of self-reported symptoms to predict potential COVID-19. *Nat. Med.* **26**, 1037–1040 (2020).
- Eliezer, M. et al. Sudden and complete olfactory loss function as a possible symptom of COVID-19. *JAMA Otolaryngol. Neck Surg.* **146**, 674–675 (2020).
- Gautier, J.-F. & Ravussin, Y. A new symptom of COVID-19: loss of taste and smell. *Obesity* **28**, 848–848 (2020).
- Pellegrino, R. et al. Corona viruses and the chemical senses: past, present, and future. *Chem. Senses* **45**, 415–422 (2020).
- Bagheri, S. H. et al. Coincidence of COVID-19 epidemic and olfactory dysfunction outbreak in Iran. *Med. J. Islam. Repub. Iran. MJIRI* **34**, 446–452 (2020).
- Beltrán-Corbellini, Á. et al. Acute-onset smell and taste disorders in the context of Covid-19: a pilot multicenter PCR-based case-control study. *Eur. J. Neurol.* **27**, 1738–1741 (2020).
- Bénézit, F. et al. Utility of hyposmia and hypogeusia for the diagnosis of COVID-19. *Lancet Infect. Dis.* **20**, 1014–1015 (2020).
- Giacomelli, A. et al. Self-reported olfactory and taste disorders in SARS-CoV-2 patients: a cross-sectional study. *Clin. Infect. Dis.* **71**, 889–890 (2020).
- Haehner, A., Drafi, J., Dräger, S., de With, K. & Hummel, T. Predictive value of sudden olfactory loss in the diagnosis of COVID-19. *ORL* **82**, 175–180 (2020).
- Yan, C. H., Faraji, F., Prajapati, D. P., Ostrander, B. T. & DeConde, A. S. Self-reported olfactory loss associates with outpatient clinical course in Covid-19. *Int. Forum Allergy Rhinol.* **10**, 821–831 (2020).
- Salje, H. et al. Estimating the burden of SARS-CoV-2 in France. *Science* **369**, 208–211 (2020).
- Levinson, R. et al. Anosmia and dysgeusia in patients with mild SARS-CoV-2 infection. Preprint at, <https://doi.org/10.1101/2020.04.11.20055483> (2020).
- Klopfenstein, T. et al. Features of anosmia in COVID-19. *Med. Mal. Infect.* **50**, 436–439 (2020).
- Lechien, J. R. et al. Olfactory and gustatory dysfunctions as a clinical presentation of mild-to-moderate forms of the coronavirus disease (COVID-19): a multicenter European study. *Eur. Arch. Otorhinolaryngol.* **277**, 2251–2261 (2020).
- Moein, S. T. et al. Smell dysfunction: a biomarker for COVID-19. *Int. Forum Allergy Rhinol.* **10**, 944–950 (2020).
- Lotsch, J. & Hummel, T. A data science-based analysis of seasonal patterns in outpatient presentations due to olfactory dysfunction. *Rhinology* **58**, 151–157 (2020).
- Potter, M. R., Chen, J. H., Lobban, N.-S. & Doty, R. L. Olfactory dysfunction from acute upper respiratory infections: relationship to season of onset. *Int. Forum Allergy Rhinol.* **10**, 706–712 (2020).
- Kang, J. W., Lee, Y. C., Han, K., Kim, S. W. & Lee, K. H. Epidemiology of anosmia in South Korea: a nationwide population-based study. *Sci. Rep.* **10**, 3717 (2020).
- Karni, N. et al. Self-rated smell ability enables highly specific predictors of COVID-19 status: a case control study in Israel. Preprint at <https://doi.org/10.1101/2020.07.30.20164327> (2020).
- Iravani, B. et al. Relationship between odor intensity estimates and COVID-19 prevalence prediction in a Swedish Population. *Chem. Senses* **45**, 449–456 (2020).
- Parma, V. et al. More than smell—COVID-19 is associated with severe impairment of smell, taste, and chemesthesis. *Chem. Senses* bjaa041. <https://doi.org/10.1093/chemse/bjaa041> (2020).

Acknowledgements

This work was supported financially by EXTREM-O (CNRS MITI), CONFINEZ2 (CNES), the CORODORAT grant (IDEX-Lyon-Université de Lyon), by the French government, through the UCAJEDI "Investments in the Future" project managed by the ANR grant No. ANR-15-IDEX-01. The authors thank Sébastien Buthion, Conceição Silva, Brigitte Perucca, Clément Blondel, Stéphanie Younès, François Maginot, and their teams from the CNRS communication units, as well as the Global Consortium for Chemosensory Research (GCCR). Deployment of the GCCR survey in multiple languages by the Pennsylvania State University (Penn State) was partially supported by funds from James and Helen Zallie and the Penn State Sensory Evaluation Center.

Author contributions

Conceived, designed the study, and wrote the paper: D.P., V.P.-L., C.F., J.G., M.B. Data acquisition and curation: D.P., V.P., M.M., M.M., E.B., O.A., J.K., M.H., J.P., R.D., C.D., S.S., E.M., M.C.F., K.W.C., E.A.S., T.T.D., A.J.B., V.P., J.E.H., T.L., C.F., J.G., M.B. Performed analysis: D.P., V.P.-L., C.F., J.G., M.B., M.M.a., M.M.o., E.B. Edited and approved the final manuscript: D.P., V.P., M.M., M.M., E.B., O.A., J.K., M.H., J.P., R.D., C.D., S.S., E.M., M.C.F., K.W.C., E.A.S., T.T.D., A.J.B., V.P., J.E.H., T.L., C.F., J.G., M.B.

Competing interests

The authors declare no competing interests.

Additional information

Supplementary information is available for this paper at <https://doi.org/10.1038/s41467-020-18963-y>.

Correspondence and requests for materials should be addressed to D.P., J.G. or M.B.

Peer review information *Nature Communications* thanks Micael Widerstrom and the other, anonymous reviewer(s) for their contribution to the peer review of this work. Peer review reports are available.

Reprints and permission information is available at <http://www.nature.com/reprints>

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this license, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2020

**Annexe 4 : « Résultats supplémentaires de l'Étude 1 »
(publié)**

Supplementary file of the paper: Using subgroup discovery to relate odor pleasantness and intensity to peripheral nervous system reactions

Maëlle Moranges, Marc Plantevit, and Moustafa Bensafi



FIGURE S1. VISUALIZATION OF THE DISCRETIZATION OF EACH SUBJECT FOR EACH PERCEPTUAL DIMENSION

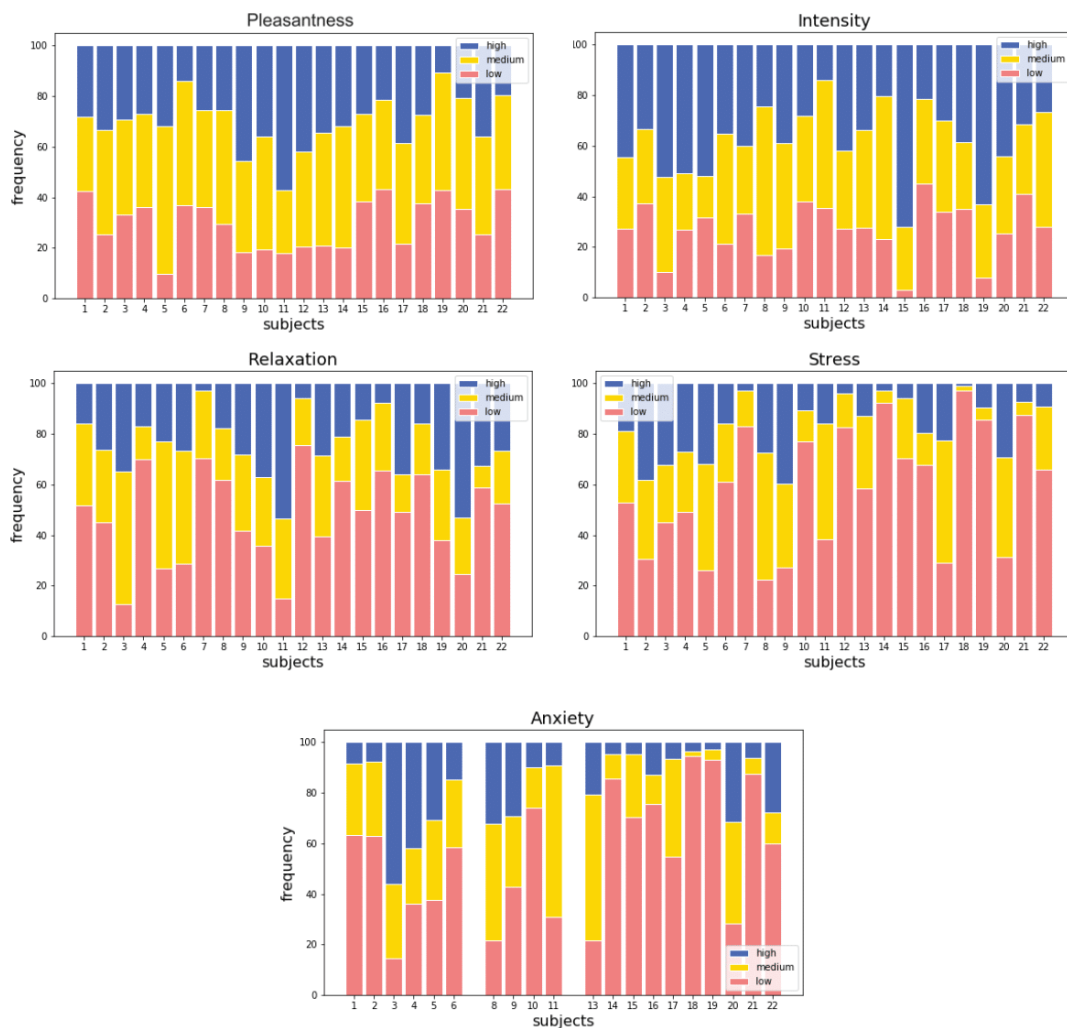


TABLE S1. RESULTS FOR PLEASANTNESS

Target	Rank	Pattern	Support_size_ratio	Informedness	Confidence interval of 10000 random sample	Number of individuals in the pattern
Unpleasant	TOP1	SC_Latency \in [0.5, 3.37109375]	0.408	0.1508	[-0.0439;0.045]	15/22
	TOP2	SC_Amplitude \in [0.01550342771861657, 1.089544904930541]	0.457	0.1407	[-0.0454;0.0455]	15/22
	TOP3	SC_Latency \in [0.5, 2.51953125]	0.317	0.1373	[-0.0426;0.0442]	15/22
	TOP4	SC_Rise_Time \in [2.21875, 9.03125]	0.355	0.136	[-0.0417;0.0429]	17/22
	TOP5	SC_Latency \in [0.03515625, 2.75]	0.396	0.1318	[-0.0438;0.0429]	12/22
	TOP6	SC_Latency \in [1.01953125, 3.16015625]	0.322	0.1312	[-0.0423;0.0423]	14/22
	TOP7	SC_Amplitude \in [0.04238531471555262, 1.339538951968894]	0.368	0.1241	[-0.0452;0.0437]	14/22
	TOP8	SC_Latency \in [0.9296875, 2.6171875]	0.281	0.1178	[-0.0409;0.0395]	16/22
	TOP9	SC_Amplitude \in [0.07501193795499272, 1.105676856073533]	0.293	0.1177	[-0.0409;0.0416]	16/22
	TOP10	SC_Latency \in [0.02734375, 2.27734375]	0.318	0.117	[-0.0417;0.0429]	14/22
Neutral	TOP1	SC_Amplitude \in [0.0, 0.01169551394785151]	0.523	0.0873	[-0.0415;0.0432]	17/22
	TOP2	SC_Rise_Time \in [0.0, 2.3125]	0.655	0.0746	[-0.0396;0.0414]	18/22
	TOP3	SC_NSSCR \in [-3.0, 0.0]	0.584	0.0635	[-0.0415;0.0414]	15/22
	TOP4	SC_Latency \in [0.0, 1.4375]	0.682	0.0604	[-0.039;0.0402]	16/22
	TOP5	FPE [-2.666666666666666, 1.0] ARE [-1.0, 0.3333333333333335] STE [-0.10991132812495508, 0.08744440104222662] SC_Latency \in [0.0, 0.02734375] SC_Rise_Time \in [0.0, 2.5078125] SC_Amplitude \in [0.0, 0.07431129858868246] SC_NSSCR \in [0.0, 2.0]	0.24	0.0574	[-0.0365;0.0353]	17/22
	TOP6	STE [-0.11054023437549532, -0.02829414062518154]	0.303	0.0556	[-0.0401;0.0372]	14/22
	TOP7	FPE [-1.0, 1.333333333333334] ARE [-1.666666666666667, 0.3333333333333335] STE [-0.09862083333355898, 0.010239062499742602] SC_Latency \in [0.0, 1.95703125] SC_Rise_Time \in [0.0, 1.56640625] SC_Amplitude \in [0.0, 0.1141159435184038] SC_NSSCR \in [-1.0, 0.0]	0.153	0.0556	[-0.0309;0.0298]	17/22
	TOP8	FPE [-1.333333333333334, 1.333333333333334] ARE [-1.666666666666667, 0.666666666666665] STE [-0.2127966145836453, 0.06015859375053623] SC_Latency \in [0.0, 0.0] SC_Rise_Time \in [0.0, 0.0] SC_Amplitude \in [0.0, 0.0] SC_NSSCR \in [0.0, 3.0]	0.279	0.055	[-0.0371;0.0366]	15/22
	TOP9	FPE [-1.0, 0.666666666666661] ARE [-1.333333333333335, 0.333333333333335] STE [-0.09862083333355898, 0.06540768229093175] SC_Latency \in [0.0, 0.0] SC_Rise_Time \in [0.0, 0.0] SC_Amplitude \in [0.0, 0.0] SC_NSSCR \in [-1.0, 0.0]	0.154	0.0541	[-0.0288;0.0302]	17/22
	TOP10	FPE [-1.0, 1.333333333333334] ARE [-0.666666666666667, 0.3333333333333335] STE [-0.09862083333355898, 0.05217057291650207] SC_Latency \in [0.0, 2.11328125] SC_Rise_Time \in [0.0, 3.59765625] SC_Amplitude \in [0.0, 0.2271654699310155] SC_NSSCR \in [-1.0, 0.0]	0.154	0.0534	[-0.0313;0.0313]	17/22
Pleasant	TOP1	SC_Amplitude \in [0.0, 0.1427779809244993]	0.793	0.0707	[-0.0364;0.0371]	14/22
	TOP2	SC_Amplitude \in [0.0, 0.03876643583032213]	0.625	0.0675	[-0.0439;0.0444]	9/22
	TOP3	FPE [-0.666666666666661, 0.0]	0.426	0.056	[-0.0449;0.0455]	15/22
	TOP4	FPE [-1.333333333333334, 0.666666666666661] ARE [-1.0, 1.666666666666667] STE [-0.16654843750031034, 0.014854687498662145] SC_Latency \in [0.0, 0.07421875] SC_Rise_Time \in [0.0, 4.35546875]	0.248	0.0553	[-0.0392;0.0385]	12/22

		SC_Amplitude \in [0.0, 0.691013917466063] SC_NSSCR \in [0.0, 3.0]				
TOP5		FP \in [-1.666666666666666, 0.6666666666666661] AR \in [-1.333333333333333, 0.0] ST \in [-0.12490690104154822, 0.1424990885412427] SC_Latency \in [0.0, 0.0] SC_Rise_Time \in [0.0, 0.0] SC_Amplitude \in [0.0, 0.0] SC_NSSCR \in [-3.0, 2.0]	0.222	0.0548	[-0.0376;0.038]	13/22
TOP6		FP \in [-0.6666666666666661, 0.6666666666666661]	0.646	0.0531	[-0.0436;0.0426]	14/22
TOP7		ST \in [-0.011125911458155002, 0.006108203124671974]	0.234	0.0525	[-0.0379;0.0378]	13/22
TOP8		SC_Latency \in [0.0, 0.4921875]	0.556	0.0519	[-0.0448;0.0435]	8/22
TOP9		SC_Rise_Time \in [0.0, 2.60546875]	0.687	0.0509	[-0.0415;0.0425]	13/22
TOP10		ST \in [-0.021952343750289316, 0.012486979166283163]	0.348	0.0503	[-0.0421;0.0419]	12/22

TABLE S2. RESULT OR INTENSITY

Target	Rank	pattern	support_size_ratio	informedness	Confidence interval of 10000 random sample	Number of individuals in the pattern
Not at all intense	TOP1	SC_NSSCR \in [-2.0, 0.0]	0.581	0.0628	[-0.0452;0.0471]	17/22
	TOP2	ST \in [-0.01067591145783453, 0.048491276041183085]	0.359	0.0628	[-0.0452;0.0448]	16/22
	TOP3	AR \in [-1.3333333333333335, -0.3333333333333326]	0.42	0.0603	[-0.0454;0.0446]	17/22
	TOP4	ST \in [-0.07068932291739571, 0.03647708333257427]	0.69	0.0602	[-0.0433;0.0422]	15/22
	TOP5	ST \in [-0.011719010418882903, 0.020387109374670587]	0.296	0.0565	[-0.0425;0.043]	15/22
	TOP6	ST \in [-0.1785326822918094, 0.058744531250326304]	0.885	0.0561	[-0.0294;0.0291]	16/22
	TOP7	ST \in [-0.007886979168432617, 0.03512265625053601]	0.287	0.0557	[-0.0411;0.0422]	15/22
	TOP8	SC_NSSCR \in [0.0, 0.0]	0.408	0.0535	[-0.0454;0.0468]	13/22
	TOP9	AR \in [-1.3333333333333335, 0.6666666666666667]	0.847	0.0512	[-0.0343;0.0332]	15/22
	TOP10	ST \in [-0.004655338543294363, 0.059873307292107825]	0.29	0.0508	[-0.0414;0.0418]	16/22
Midly intense	TOP1	SC_Latency \in [0.0, 0.48046875]	0.553	0.088	[-0.0443;0.0433]	16/22
	TOP2	SC_Amplitude \in [0.0, 0.07361034640270958]	0.701	0.0769	[-0.0398;0.04]	15/22
	TOP3	SC_Rise_Time \in [0.0, 2.41015625]	0.665	0.0749	[-0.0399;0.0418]	14/22
	TOP4	SC_Latency \in [0.0, 1.53125]	0.695	0.0618	[-0.0414;0.0384]	14/22
	TOP5	FP \in [-1.0, 1.333333333333334] AR \in [-1.666666666666667, 0.3333333333333335] ST \in [-0.100633072916672, 0.010239062499742602] SC_Latency \in [0.0, 1.95703125] SC_Rise_Time \in [0.0, 1.56640625] SC_Amplitude \in [0.0, 0.1141159435184038] SC_NSSCR \in [-1.0, 2.0]	0.221	0.0577	[-0.0357;0.0363]	13/22
	TOP6	FP \in [-0.6666666666666661, 2.333333333333334] AR \in [-1.0, 2.666666666666665] ST \in [-0.18510924479184254, 0.0007335937492278788] SC_Latency \in [0.0, 0.3125] SC_Rise_Time \in [0.0, 1.625] SC_Amplitude \in [0.0, 0.05957652442768291] SC_NSSCR \in [-1.0, 1.0]	0.222	0.053	[-0.0345;0.0355]	13/22
	TOP7	FP \in [-0.3333333333333339, 1.0] AR \in [-1.0, 0.0] ST \in [-0.09231822916678567, 0.05891718750000052] SC_Latency \in [0.0, 0.0] SC_Rise_Time \in [0.0, 0.0] SC_Amplitude \in [0.0, 0.0] SC_NSSCR \in [-1.0, 3.0]	0.138	0.0525	[-0.0292;0.0292]	15/22
	TOP8	FP \in [-1.0, 1.666666666666666] AR \in [-1.0, 0.0] ST \in [-0.07068932291739571, 0.13076848958374399] SC_Latency \in [0.0, 0.453125] SC_Rise_Time \in [0.0, 1.16015625]	0.198	0.0522	[-0.0335;0.0347]	16/22

		SC_Amplitude∈ [0.0, 0.06123546517595146] SC_NSSCR∈ [-2.0, 4.0]				
	TOP9	FPE [-2.33333333333334, 0.666666666666661] ARE [-0.666666666666665, 1.333333333333335] STE [-0.08805950520766004, 0.11037643229136052] SC_Latency∈ [0.0, 0.11328125] SC_Rise_Time∈ [0.0, 1.2421875] SC_Amplitude∈ [0.0, 0.01469741673709613] SC_NSSCR∈ [-2.0, 0.0]	0.208	0.0521	[-0.0354;0.0346]	15/22
	TOP10	FPE [-1.0, 1.0] ARE [-1.0, 0.333333333333326] STE [-0.09231822916678567, 0.07767447916669212] SC_Latency∈ [0.0, 2.3515625] SC_Rise_Time∈ [0.0, 1.71484375] SC_Amplitude∈ [0.0, 0.06767383559014917] SC_NSSCR∈ [-1.0, 1.0]	0.175	0.0518	[-0.0318;0.0324]	16/22
Very intense	TOP1	SC_Latency∈ [0.02734375, 3.42578125]	0.458	0.1229	[-0.042;0.0423]	15/22
	TOP2	SC_Amplitude∈ [0.02274101545841711, 0.552357904828892]	0.39	0.1094	[-0.0424;0.0419]	15/22
	TOP3	SC_Rise_Time∈ [1.95703125, 8.296875]	0.391	0.1023	[-0.0401;0.0423]	14/22
	TOP4	SC_Latency∈ [0.58203125, 2.93359375]	0.362	0.1008	[-0.0398;0.0408]	16/22
	TOP5	SC_Latency∈ [0.046875, 2.53515625]	0.36	0.0943	[-0.0407;0.0399]	15/22
	TOP6	SC_Latency∈ [0.953125, 3.47265625]	0.361	0.094	[-0.0409;0.0415]	15/22
	TOP7	SC_Amplitude∈ [0.01111802570735487, 0.3491161441646504]	0.399	0.0925	[-0.0425;0.0419]	15/22
	TOP8	SC_Amplitude∈ [0.07410586586195356, 0.9904480187709754]	0.295	0.09	[-0.0393;0.0394]	12/22
	TOP9	SC_NSSCR∈ [1.0, 5.0]	0.412	0.0893	[-0.0419;0.0424]	19/22
	TOP10	SC_Amplitude∈ [0.02065301884808202, 0.2598925037991053]	0.316	0.0871	[-0.0385;0.0402]	15/22

TABLE S3. RESULTS FOR RELAXATION

Target	Rank	pattern	support_size_ratio	informedness	Confidence interval of 10000 random sample	Number of individuals in the pattern
Not at all relaxing	TOP1	ARE [-0.333333333333337, 0.333333333333326]	0.313	0.0745	[-0.0391;0.039]	13/22
	TOP2	SC_Rise_Time∈ [1.82421875, 10.4921875]	0.412	0.0736	[-0.04;0.0399]	10/22
	TOP3	ARE [-0.333333333333337, 0.666666666666667]	0.529	0.0731	[-0.0423;0.0411]	11/22
	TOP4	STE [-0.08816210937516189, 0.019982552083867233]	0.687	0.0719	[-0.0382;0.0364]	16/22
	TOP5	SC_Latency∈ [1.6484375, 3.75]	0.28	0.0716	[-0.0367;0.036]	15/22
	TOP6	SC_Latency∈ [1.078125, 3.44140625]	0.338	0.0704	[-0.0397;0.0385]	13/22
	TOP7	STE [-0.0464398437500968, 0.027019531251596618]	0.561	0.0688	[-0.0413;0.0404]	15/22
	TOP8	SC_Rise_Time∈ [2.4921875, 10.4921875]	0.326	0.068	[-0.0403;0.0378]	10/22
	TOP9	STE [-0.04875885416600667, 0.002023307291945997]	0.47	0.0662	[-0.0421;0.0396]	14/22
	TOP10	STE [-0.09410130208298284, -0.0027028645820372788]	0.552	0.0651	[-0.0396;0.0403]	14/22
Midly relaxing	TOP1	ARE [-2.333333333333335, -0.333333333333335]	0.34	0.0875	[-0.0434;0.0431]	16/22
	TOP2	SC_Rise_Time∈ [0.0, 2.21875]	0.645	0.0708	[-0.0446;0.0441]	11/22
	TOP3	ARE [-2.333333333333335, -0.666666666666665]	0.271	0.0635	[-0.0407;0.0392]	12/22
	TOP4	ARE [-2.333333333333335, -0.333333333333326]	0.441	0.0587	[-0.0456;0.0454]	14/22
	TOP5	ARE [-2.333333333333335, -0.666666666666667]	0.199	0.0537	[-0.035;0.0359]	14/22
	TOP6	SC_Amplitude∈ [0.0, 0.3212598751120668]	0.906	0.0517	[-0.0281;0.0273]	12/22
	TOP7	ARE [-2.666666666666667, -0.666666666666667]	0.154	0.0516	[-0.0327;0.0339]	13/22
	TOP8	FPE [-6.666666666666666, -0.666666666666661]	0.353	0.0514	[-0.0439;0.0426]	14/22
	TOP9	SC_Amplitude∈ [0.0, 0.07676639593778822]	0.708	0.0507	[-0.0425;0.0418]	13/22

	TOP10	SC_Latency€ [0.0, 0.03515625]	0.514	0.0495	[-0.0458;0.0451]	15/22
Very relaxing	TOP1	SC_Latency€ [0.0, 1.6953125]	0.722	0.0623	[-0.0428;0.0436]	15/22
	TOP2	ST€ [-0.0014286458342311903, 0.1000091145905898]	0.283	0.0555	[-0.0425;0.0416]	14/22
	TOP3	ST€ [0.004131770832284332, 0.0952856770834245]	0.222	0.0521	[-0.039;0.0404]	14/22
	TOP4	FPE [-0.6666666666666661, 0.0]	0.426	0.0497	[-0.0461;0.0473]	13/22
	TOP5	ST€ [0.002023307291945997, 0.20339739583171124]	0.278	0.0474	[-0.0413;0.0428]	11/22
	TOP6	AR€ [0.3333333333333326, 3.33333333333333]	0.371	0.0467	[-0.0467;0.0444]	14/22
	TOP7	ST€ [0.023384895834052344, 0.10104778645828105]	0.145	0.0443	[-0.0327;0.0327]	11/22
	TOP8	FPE [-0.6666666666666661, 2.0]	0.755	0.0435	[-0.0405;0.0412]	12/22
	TOP9	AR€ [-0.333333333333335, 3.0]	0.728	0.0429	[-0.0411;0.0406]	15/22
	TOP10	SC_Latency€ [0.02734375, 1.765625]	0.223	0.0421	[-0.0372;0.0398]	13/22

TABLE S4. RESULTS FOR STRESS

Target	Rank	pattern	support_size_ratio	informedness	Confidence interval of 10000 random sample	Number of individuals in the pattern
Not at all stressful	TOP1	SC_Rise_Time€ [1.1015625, 6.703125]	0.462	0.0974	[-0.0424;0.0411]	9/22
	TOP2	SC_Rise_Time€ [2.01171875, 7.765625]	0.379	0.0899	[-0.0408;0.0409]	8/22
	TOP3	SC_Rise_Time€ [1.7421875, 5.16015625]	0.367	0.083	[-0.0404;0.0395]	8/22
	TOP4	SC_Latency€ [0.08984375, 2.296875]	0.315	0.0809	[-0.0389;0.0392]	9/22
	TOP5	SC_Latency€ [0.07421875, 2.8203125]	0.401	0.0793	[-0.0405;0.0412]	9/22
	TOP6	SC_Latency€ [0.69921875, 3.55078125]	0.399	0.0762	[-0.04;0.0399]	11/22
	TOP7	SC_Rise_Time€ [2.65625, 6.55859375]	0.291	0.0717	[-0.0373;0.0372]	7/22
	TOP8	SC_Rise_Time€ [1.2265625, 4.4140625]	0.369	0.0697	[-0.041;0.0407]	12/22
	TOP9	SC_Latency€ [0.09765625, 1.8359375]	0.231	0.0685	[-0.035;0.034]	11/22
	TOP10	SC_Amplitude€ [0.1831605675242298, 0.9118806529445909]	0.165	0.0685	[-0.0314;0.0304]	5/22
Midly stressful	TOP1	SC_Amplitude€ [0.0, 0.01213742320049771]	0.525	0.1303	[-0.0468;0.0489]	15/22
	TOP2	SC_Amplitude€ [0.0, 0.09674000864212795]	0.733	0.1166	[-0.0413;0.04]	12/22
	TOP3	FPE [-3.0, 0.6666666666666661] AR€ [-0.6666666666666667, 1.333333333333335] ST€ [-0.054363020834138354, 0.08490794270896629] SC_Latency€ [0.0, 0.28515625] SC_Rise_Time€ [0.0, 4.92578125] SC_Amplitude€ [0.0, 0.4512014489457026] SC_NSSCR€ [-1.0, 0.0]	0.198	0.1016	[-0.0373;0.0393]	15/22
	TOP4	FPE [-1.333333333333334, 0.6666666666666661] AR€ [-0.6666666666666665, 2.666666666666667] ST€ [-0.17873489583304192, 0.11037643229136052] SC_Latency€ [0.0, 2.625] SC_Rise_Time€ [0.0, 1.8515625] SC_Amplitude€ [0.0, 0.01586861084889213] SC_NSSCR€ [0.0, 1.0]	0.227	0.1012	[-0.04;0.0414]	13/22
	TOP5	SC_Latency€ [0.0, 1.05859375]	0.631	0.0909	[-0.0455;0.0478]	14/22
	TOP6	SC_Rise_Time€ [0.0, 2.296875]	0.653	0.0904	[-0.0461;0.0449]	9/22
	TOP7	FPE [-1.666666666666666, 0.6666666666666661] AR€ [-0.333333333333335, 1.333333333333335] ST€ [-0.04900976562560189, 0.06534166666688179] SC_Latency€ [0.0, 1.99609375] SC_Rise_Time€ [0.0, 2.11328125] SC_Amplitude€ [0.0, 0.1180602851636334] SC_NSSCR€ [-1.0, 0.0]	0.153	0.0901	[-0.0344;0.035]	15/22
	TOP8	FPE [-1.666666666666666, 1.333333333333334] AR€ [-1.0, 1.333333333333335] ST€ [-0.12714361979187672, 0.06853854166651985] SC_Latency€ [0.0, 1.21875] SC_Rise_Time€ [0.0, 4.015625]	0.328	0.0898	[-0.0442;0.0467]	9/22

		SC_Amplitude∈ [0.0, 0.2468485182836724] SC_NSSCR∈ [0.0, 2.0]				
	TOP9	FPE [-1.0, 0.3333333333333339] ARE [-1.333333333333333, 1.333333333333335] STE [-0.1785326822918094, 0.080793489582895] SC_Latency∈ [0.0, 0.0] SC_Rise_Time∈ [0.0, 0.0] SC_Amplitude∈ [0.0, 0.0] SC_NSSCR∈ [-1.0, 4.0]	0.274	0.087	[-0.0422;0.0416]	16/22
	TOP10	FPE [-3.0, 0.3333333333333339] ARE [-0.333333333333335, 1.333333333333335] STE [-0.06522408854219819, 0.08490794270896629] SC_Latency∈ [0.0, 0.0] SC_Rise_Time∈ [0.0, 0.0] SC_Amplitude∈ [0.0, 0.0] SC_NSSCR∈ [-1.0, 2.0]	0.213	0.084	[-0.038;0.0386]	12/22
Very stressful	TOP1	ARE [-0.666666666666665, 0.666666666666665]	0.63	0.0824	[-0.052;0.0518]	13/22
	TOP2	SC_NSSCR∈ [1.0, 5.0]	0.412	0.078	[-0.0534;0.0536]	14/22
	TOP3	SC_NSSCR∈ [2.0, 5.0]	0.178	0.0746	[-0.0415;0.041]	11/22
	TOP4	FPE [-5.333333333333334, -1.0]	0.235	0.0726	[-0.0466;0.0451]	11/22
	TOP5	FPE [-5.0, -0.666666666666661]	0.351	0.0726	[-0.0527;0.0512]	12/22
	TOP6	SC_Amplitude∈ [0.0826669580983157, 0.3766110044964068]	0.211	0.0624	[-0.0445;0.0441]	14/22
	TOP7	FPE [-5.0, -1.333333333333334]	0.155	0.0622	[-0.0386;0.0408]	11/22
	TOP8	SC_Amplitude∈ [0.05738522482499298, 0.3471000276938734]	0.253	0.0573	[-0.0466;0.0481]	13/22
	TOP9	ARE [0.3333333333333304, 0.666666666666665]	0.206	0.0561	[-0.0417;0.0469]	13/22
	TOP10	ARE [-0.3333333333333304, 0.666666666666665]	0.387	0.0556	[-0.0514;0.0525]	15/22

TABLE S5. RESULTS FOR ANXIETY

Target	Rank	pattern	support_size_ratio	informedness	Confidence interval of 10000 random sample	Number of individuals in the pattern
Not at all anxiety-inducing	TOP1	SC_Latency∈ [0.05078125, 1.8515625]	0.24	0.0858	[-0.0364;0.0365]	10/20
	TOP2	SC_Rise_Time∈ [1.90625, 6.46875]	0.367	0.0797	[-0.0406;0.0422]	7/20
	TOP3	SC_Latency∈ [0.03515625, 2.1953125]	0.302	0.0781	[-0.0402;0.0407]	8/20
	TOP4	SC_Amplitude∈ [0.2055990804015031, 0.9349887955018843]	0.159	0.0778	[-0.0326;0.0324]	4/20
	TOP5	SC_Latency∈ [0.06640625, 3.00390625]	0.405	0.0773	[-0.0429;0.0418]	8/20
	TOP6	SC_Amplitude∈ [0.2562295491428905, 1.186667060070035]	0.128	0.0747	[-0.0298;0.0294]	5/20
	TOP7	SC_Latency∈ [0.09765625, 1.47265625]	0.169	0.0726	[-0.0319;0.0332]	11/20
	TOP8	SC_Rise_Time∈ [1.640625, 4.9296875]	0.352	0.0701	[-0.0402;0.0406]	9/20
	TOP9	STE [-0.15181809895845433, -0.032432812499749275]	0.335	0.067	[-0.0395;0.0413]	14/20
	TOP10	SC_Latency∈ [0.22265625, 1.6953125]	0.189	0.0633	[-0.0353;0.0337]	10/20
Midly anxiety-inducing	TOP1	SC_Amplitude∈ [0.0, 0.01882392073889472]	0.566	0.1494	[-0.0484;0.0467]	11/20
	TOP2	SC_Rise_Time∈ [0.0, 2.5859375]	0.704	0.1172	[-0.0455;0.0446]	13/20
	TOP3	FPE [-0.666666666666661, 0.666666666666661] ARE [-1.0, 1.0] STE [-0.0874880208329678, 0.06421536458330479] SC_Latency∈ [0.0, 0.0] SC_Rise_Time∈ [0.0, 0.0] SC_Amplitude∈ [0.0, 0.0] SC_NSSCR∈ [-3.0, 0.0]	0.162	0.1075	[-0.0352;0.0374]	13/20
	TOP4	FPE [-1.666666666666666, 1.333333333333334] ARE [-1.0, 1.333333333333335] STE [-0.12714361979187672, 0.06853854166651985] SC_Latency∈ [0.0, 1.21875] SC_Rise_Time∈ [0.0, 4.015625] SC_Amplitude∈ [0.0, 0.2468485182836724]	0.336	0.1071	[-0.0457;0.047]	10/20

		SC_NSSCR€ [0.0, 2.0]				
	TOP5	FP€ [-0.6666666666666661, 1.3333333333333334] AR€ [-1.0, 1.3333333333333335] ST€ [-0.12714361979187672, 0.13837760416701173] SC_Latency€ [0.0, 0.0] SC_Rise_Time€ [0.0, 0.0] SC_Amplitude€ [0.0, 0.0] SC_NSSCR€ [-1.0, 0.0]	0.204	0.1013	[-0.039;0.0387]	12/20
	TOP6	SC_Amplitude€ [0.0, 0.2034830910078747]	0.835	0.0986	[-0.0366;0.0361]	7/20
	TOP7	FP€ [-1.0, 0.3333333333333339] AR€ [-1.6666666666666665, 1.0] ST€ [-0.04861627604165619, 0.18732109374997208] SC_Latency€ [0.0, 1.109375] SC_Rise_Time€ [0.0, 2.03515625] SC_Amplitude€ [0.0, 0.01142487016627514] SC_NSSCR€ [0.0, 2.0]	0.2	0.0984	[-0.0393;0.0383]	11/20
	TOP8	FP€ [-1.0, 0.6666666666666661] AR€ [-1.0, 0.6666666666666667] ST€ [-0.0874880208329678, 0.07219023437468763] SC_Latency€ [0.0, 0.0] SC_Rise_Time€ [0.0, 0.0] SC_Amplitude€ [0.0, 0.0] SC_NSSCR€ [-3.0, 1.0]	0.218	0.0971	[-0.0406;0.0395]	14/20
	TOP9	FP€ [-0.6666666666666661, 0.3333333333333339] AR€ [-2.333333333333333, 1.0] ST€ [-0.07854934895831533, 0.10104778645828105] SC_Latency€ [0.0, 1.12109375] SC_Rise_Time€ [0.0, 0.953125] SC_Amplitude€ [0.0, 0.03220451747687214] SC_NSSCR€ [-1.0, 0.0]	0.149	0.0935	[-0.0342;0.0359]	11/20
	TOP10	FP€ [-0.6666666666666661, 1.3333333333333334] AR€ [-1.0, 1.0] ST€ [-0.025031510417761638, 0.10104778645828105] SC_Latency€ [0.0, 1.12109375] SC_Rise_Time€ [0.0, 0.953125] SC_Amplitude€ [0.0, 0.03220451747687214] SC_NSSCR€ [-1.0, 4.0]	0.202	0.0932	[-0.0395;0.0381]	14/20
Very anxiety-inducing	TOP1	SC_Amplitude€ [0.02476587272735109, 0.2731548512239026]	0.29	0.1236	[-0.0501;0.0515]	15/20
	TOP2	SC_Amplitude€ [0.01787925554868752, 0.4702459362744076]	0.38	0.1183	[-0.0554;0.0528]	15/20
	TOP3	SC_Amplitude€ [0.01000308151151996, 0.2402429916840008]	0.328	0.1099	[-0.0539;0.0509]	15/20
	TOP4	SC_Latency€ [0.95703125, 3.72265625]	0.357	0.101	[-0.053;0.0518]	15/20
	TOP5	SC_Latency€ [1.4609375, 3.75]	0.287	0.0976	[-0.0498;0.0518]	15/20
	TOP6	SC_Amplitude€ [0.05000356660674665, 0.2451090614648795]	0.215	0.0973	[-0.0436;0.0481]	16/20
	TOP7	AR€ [-0.6666666666666665, 0.6666666666666665]	0.628	0.0971	[-0.0536;0.0545]	12/20
	TOP8	SC_Latency€ [0.1640625, 3.8984375]	0.454	0.0966	[-0.0541;0.0573]	17/20
	TOP9	SC_Amplitude€ [0.01944185226604889, 0.173125589700255]	0.243	0.0961	[-0.048;0.047]	14/20
	TOP10	SC_Latency€ [1.81640625, 3.81640625]	0.231	0.094	[-0.0469;0.0482]	16/20

FIGURE S2. BOOTSTRAPPING DISTRIBUTIONS FOR THE TOP1 PATTERN OF EACH DIMENSION AND EACH CLASS

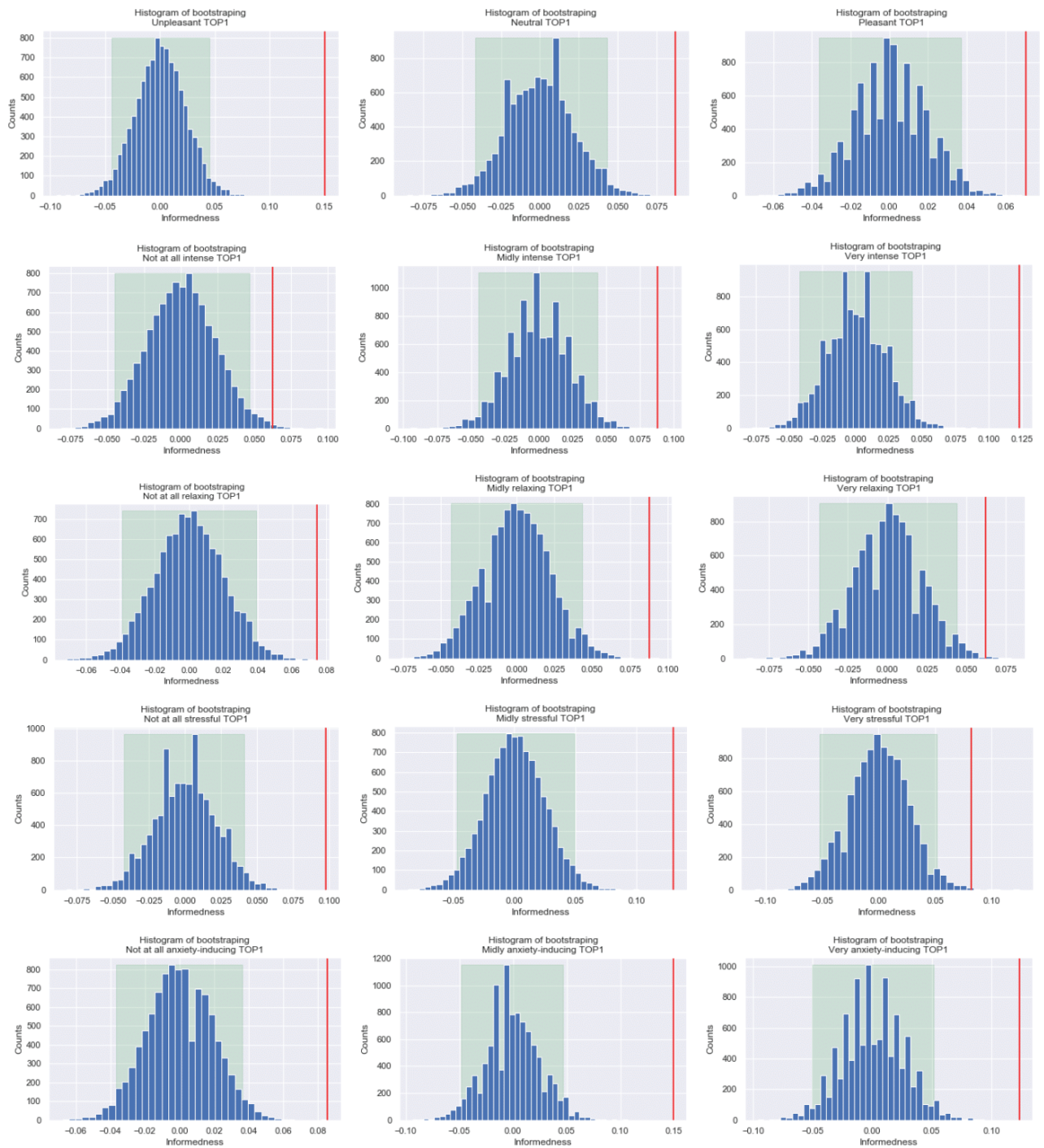


FIGURE S3. MULTIPLE RULES PATTERNS





FIGURE S4. NUMBER OF PARTICIPANTS FOR THE TOP1 PATTERN OF EACH EMOTIVE STATE

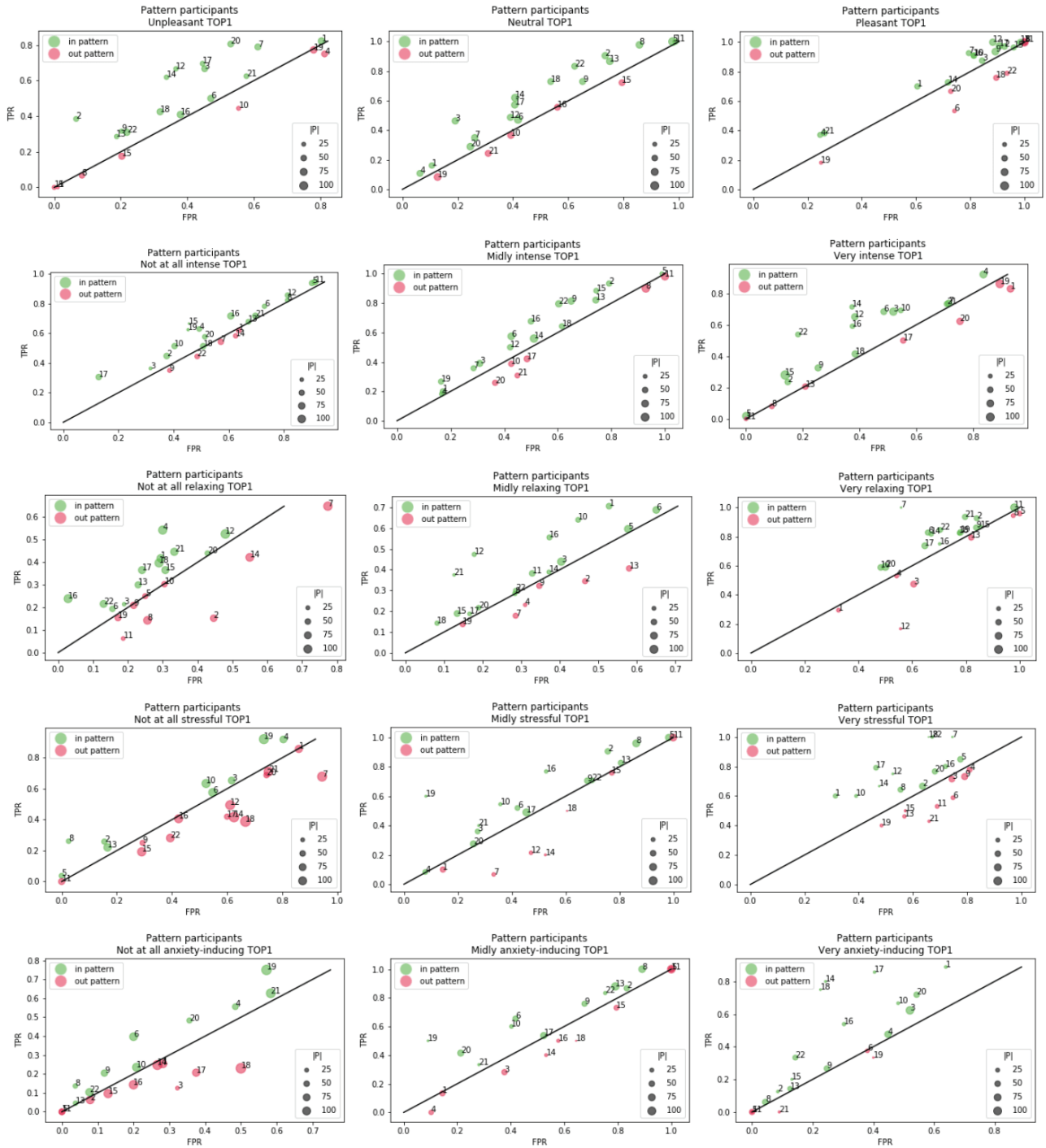
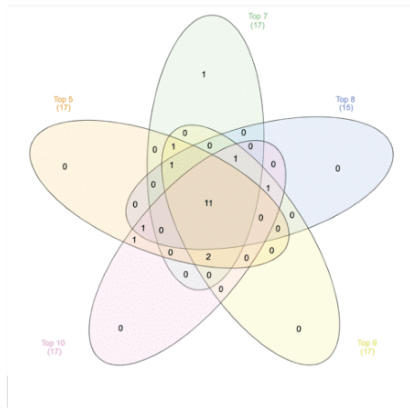


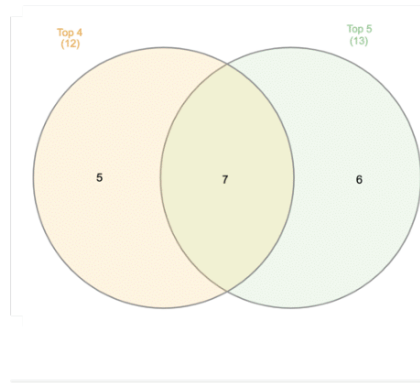
FIGURE S5. VENN DIAGRAMS OF THE OVERLAP OF INDIVIDUALS OF DIFFERENT UNIDIMENSIONAL PATTERNS FOR A SAME CATEGORY



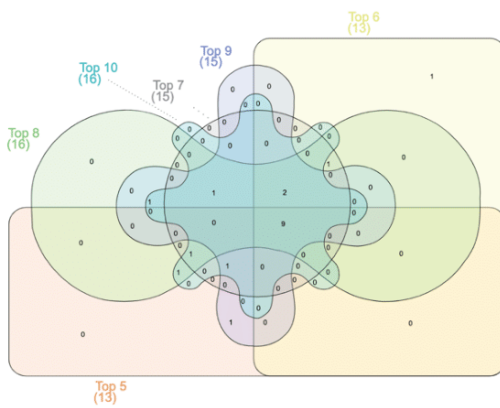
FIGURE S6. VENN DIAGRAMS OF THE OVERLAP OF INDIVIDUALS OF DIFFERENT MULTI-DIMENSIONAL PATTERNS FOR A SAME CATEGORY



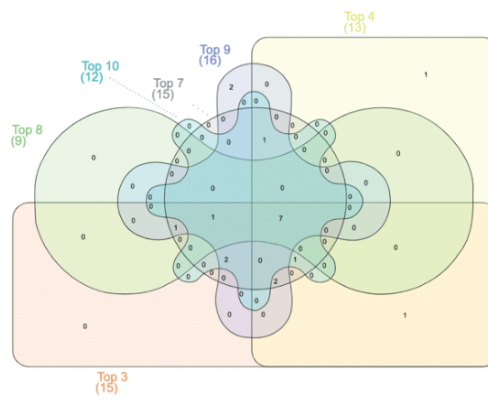
Neutral



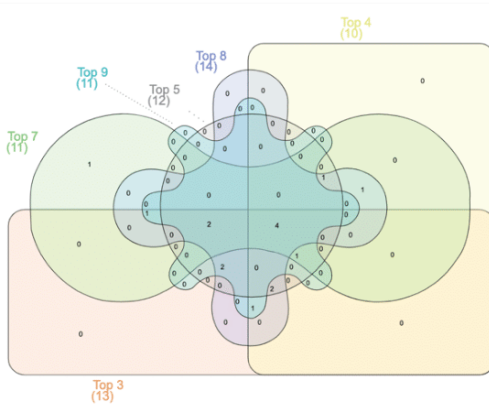
Pleasant



Midly intense



Midly stressful



Midly anxiety-inducing

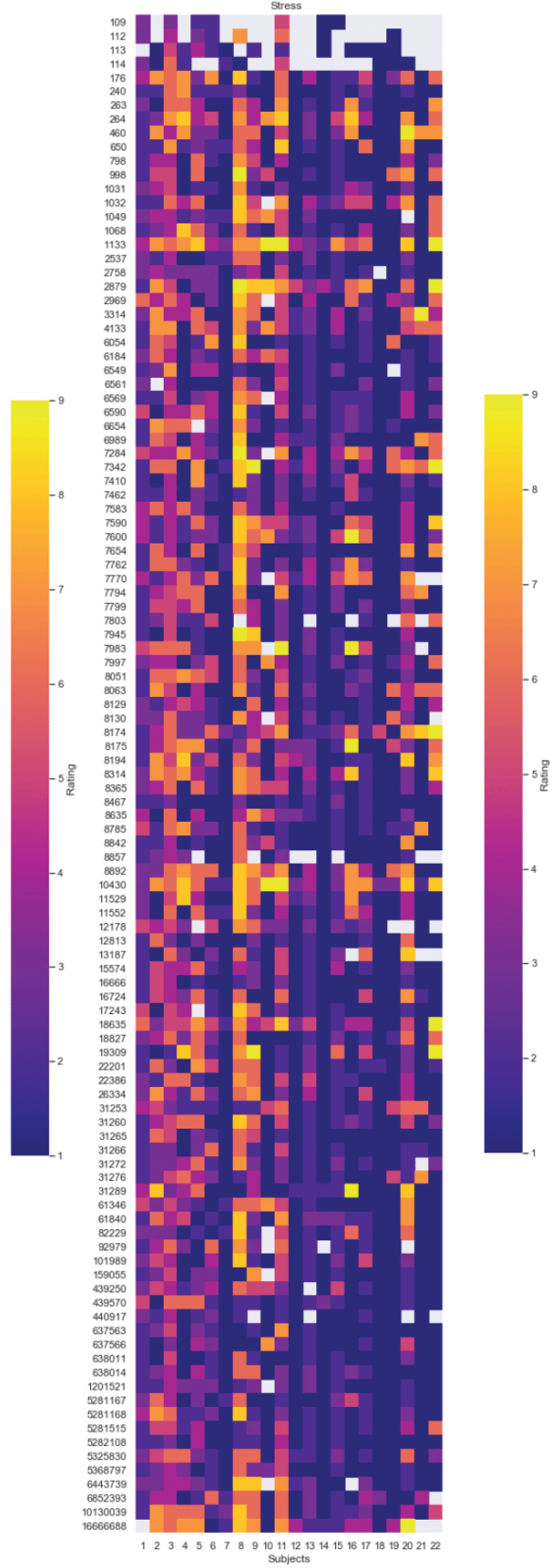
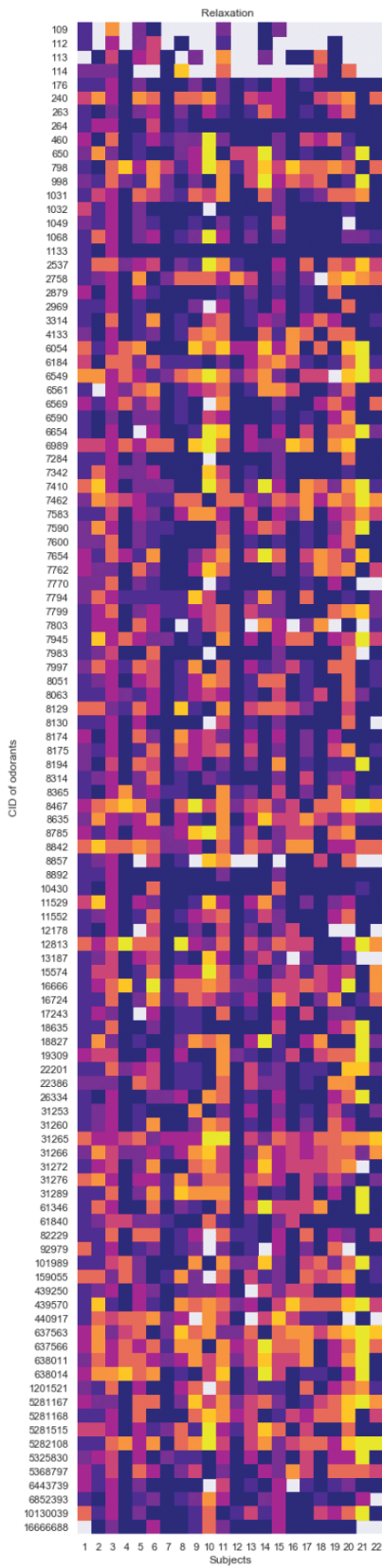
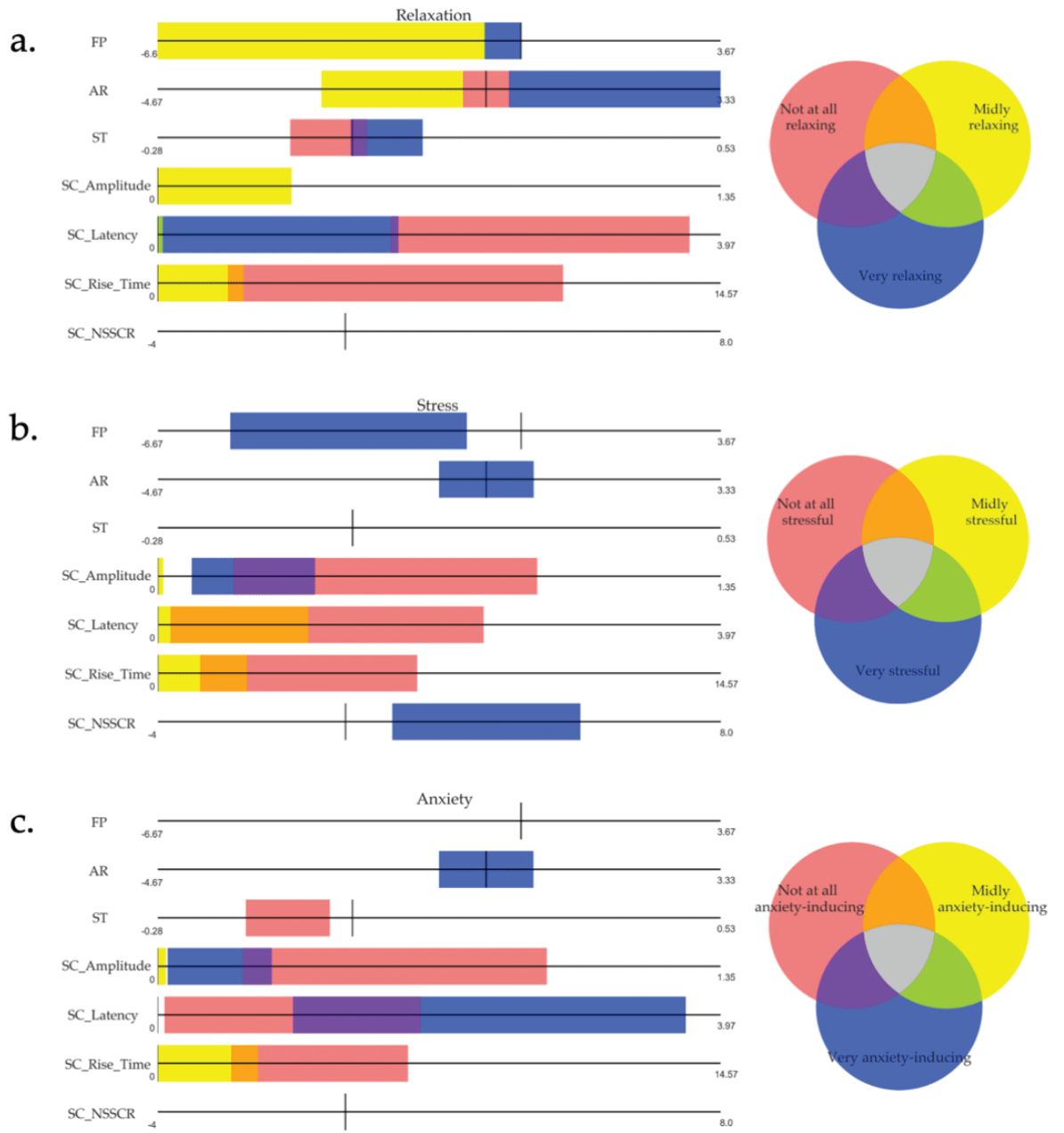


TABLE S6. MEAN 3-FOLD CROSS-VALIDATION ACCURACY (%) OF THE BEST ESTIMATORS

		DT	RF	Log. Reg.	KNN	SVM	MLP	Size of the largest class
Pleasantness	Best accuracy	39.3	39.8	39.5	34.4	40.3	41.2	40.1
	Parameters	criterion: entropy, max_depth: 3	max_depth: 3, max_features: auto, n_estimators: 10	multi_class: auto, solver: newton-cg	metric: euclidean, n_neighbors: 3, weights: distance	C: 0.1, gamma: scale, kernel: poly	activation: relu, alpha: 0.0001, hidden_layer_sizes: (50, 100, 50), learning_rate: constant, solver: sgd	
Intensity	Best accuracy	38.7	40.7	40.6	38	41.5	40.8	38.2
	Parameters	criterion: gini, max_depth: 7	max_depth: 3, max_features: auto, n_estimators: 710	multi_class: ovr, solver: saga	metric: manhattan, n_neighbors: 16, weights: uniform	C: 0.1, gamma: auto, kernel: rbf	activation: logistic, alpha: 0.0001, hidden_layer_sizes: (50, 50, 50), learning_rate: constant, solver: adamactivation: logistic, alpha: 0.0001, hidden_layer_sizes: (50, 50, 50), learning_rate: constant, solver: adam	
Relaxation	Best accuracy	45.9	47	47.1	41.8	47.1	47.1	47.1
	Parameters	criterion: entropy, max_depth: 3	max_depth: 3, max_features: auto, n_estimators: 110	multi_class: auto, solver: newton-cg	metric: manhattan, n_neighbors: 20, weights: uniform	C: 0.1, gamma: scale, kernel: linear	activation: logistic, alpha: 0.0001, hidden_layer_sizes: (50, 100, 50), learning_rate: constant, solver: adam	
Stress	Best accuracy	57.3	58	57.8	54.3	58	58.1	58.0
	Parameters	criterion: entropy, max_depth: 3	max_depth: 3, max_features: auto, n_estimators: 10	multi_class: ovr, solver: newton-cg	metric: manhattan, n_neighbors: 20, weights: uniform	C: 0.1, gamma: scale, kernel: linear	activation: identity, alpha: 0.0001, hidden_layer_sizes: (100,), learning_rate: constant, solver: sgd	
Anxiety	Best accuracy	55.3	55.4	55	50.2	55.4	55.4	55.4
	Parameters	criterion: entropy, max_depth: 3	max_depth: 3, max_features: auto, n_estimators: 110	multi_class: auto, solver: newton-cg	metric: euclidean, n_neighbors: 20, weights: uniform	C: 0.1, gamma: scale, kernel: linear	activation: logistic, alpha: 0.0001, hidden_layer_sizes: (50, 50, 50), learning_rate: constant, solver: sgd	

FIGURE S8. VISUALIZATION OF SIMPLE RULES FOR RELAXATION (A), STRESS (B) AND ANXIETY (C).



Annexe 5 : « Résultats supplémentaires de l'Étude 2 »
(en préparation)

Supplementary file of the paper:

Using exceptional attributed subgraph mining to explore interindividual variability in odor pleasantness processing in the human piriform cortex and amygdala

Maëlle Moranges, Arnaud Fournel, Marc Thévenet, Marc Plantevit, Moustafa Bensafi

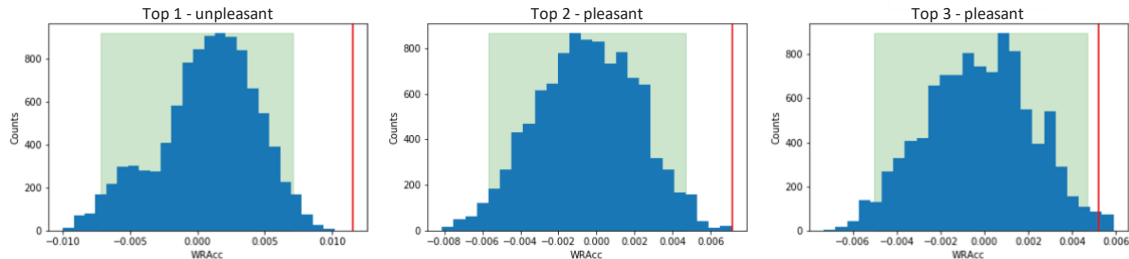


Figure S1. Bootstrapping distributions for each pattern found in the Anterior Piriform Cortex ROI. The fact that the WRAcc of the pattern (vertical bar) is above the upper limit of the confidence interval shows us that the pattern is significant.

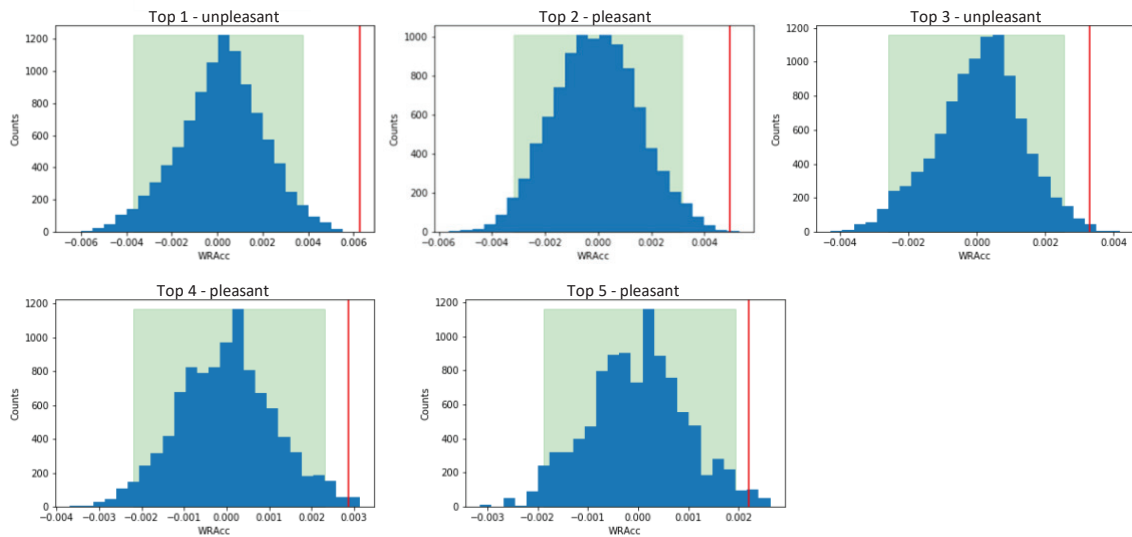


Figure S2. Bootstrapping distributions for each pattern found in the Posterior Piriform Cortex ROI. The fact that the WRAcc of the pattern (vertical bar) is above the upper limit of the confidence interval shows us that the pattern is significant.

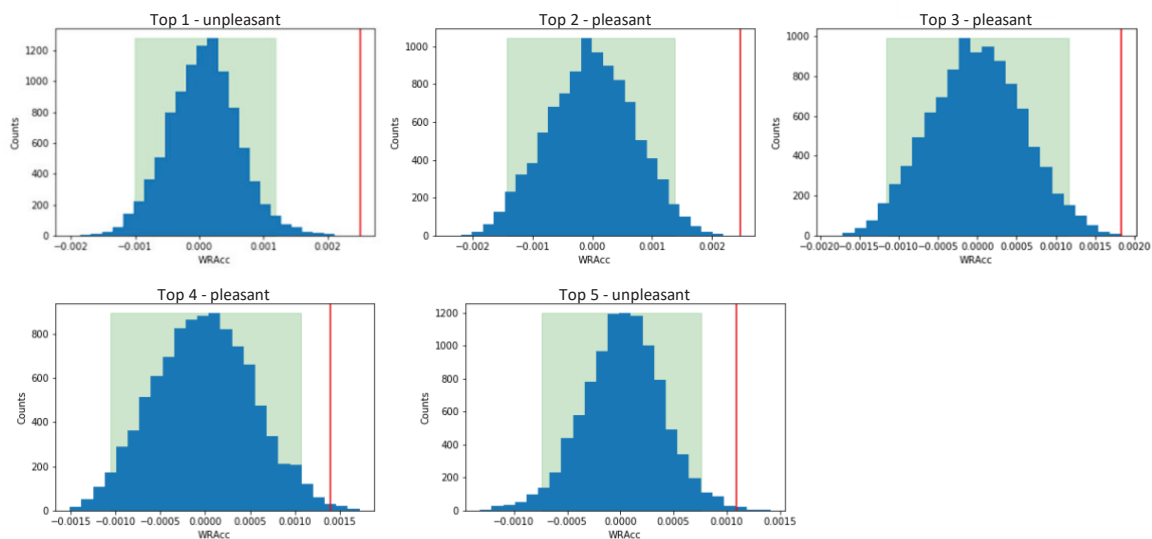


Figure S3. Bootstrapping distributions for each pattern found in the Amygdala ROI. The fact that the WRAcc of the pattern (vertical bar) is above the upper limit of the confidence interval shows us that the pattern is significant.