



HAL
open science

Are odors the best to tell memories?: cognitive and neural singularities of episodic memory evoked by odors, music, and faces

Lucile Rey

► **To cite this version:**

Lucile Rey. Are odors the best to tell memories?: cognitive and neural singularities of episodic memory evoked by odors, music, and faces. Neuroscience. Université de Lyon, 2022. English. NNT : 2022LYSE1017 . tel-03855277

HAL Id: tel-03855277

<https://theses.hal.science/tel-03855277>

Submitted on 16 Nov 2022

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.



N°d'ordre NNT :
2022LYSE1017

THESE de DOCTORAT DE L'UNIVERSITE DE LYON
opérée au sein de
l'Université Claude Bernard Lyon 1

Ecole Doctorale N° 476
Neurosciences & Cognition

Spécialité de doctorat :
Discipline : Neurosciences

Soutenue publiquement le 07/02/2022, par :
Lucile Rey

Les odeurs, les meilleures conteuses de souvenirs ?
Particularités cognitives et neuronales de la mémoire
épisodique indicée par des odeurs, musiques et visages.

Devant le jury composé de :

DELPLANQUE, Sylvain Chercheur, Univ. de Genève
GROUSSARD, Mathilde Maître de Conférence, UNICAEN
MOULIN, Christopher Prof. des Universités, UGA
PERRIN, Fabien Maître de Conférence, UCBL1
PIOLINO, Pascale Prof. des Universités, Univ. Paris Descartes

Rapporteur
Invitée
Président
Examineur
Rapporteure

PLAILLY, Jane Chargée de recherche, CNRS
TILLMANN, Barbara Directrice de Recherche, CNRS

Directrice de thèse
Co-directrice 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	Mme Céline BROCHIER
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 (ISTR)	Directeur : M. Xavier PERROT
Institut des Sciences Pharmaceutiques et Biologiques (ISBP)	Directrice : Mme Christine VINCIGUERRA

COMPOSANTES & DEPARTEMENTS DE SCIENCES & TECHNOLOGIE

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

Are odors the best to tell memories?

Cognitive and neural singularities of episodic memory evoked by odors, music, and faces.

~ ~ ~

“And yet are not these all that the illimitable capacity of my memory retains. Here also is all that is apprehended of the liberal sciences, and not yet forgotten—removed as it were into an inner place, which is not a place; nor are they the images which are retained, but the things themselves. For what is literature, what skill in disputation, whatsoever I know of all the many kinds of questions there are, is so in my memory, as that I have not taken in the image and left the thing without, or that it should have sounded and passed away like a voice imprinted on the ear by that trace, whereby it might be recorded, as though it sounded when it no longer did so; or as an odor while it passes away, and vanishes into wind, affects the sense of smell, whence it conveys the image of itself into the memory, which we realize in recollecting; or like food, which assuredly in the belly hath now no taste, and yet hath a kind of taste in the memory, or like anything that is by touching felt by the body, and which even when removed from us is imagined by the memory. For these things themselves are not put into it, but the images of them only are caught up, with a marvelous quickness, and laid up, as it were, in most wonderful garners, and wonderfully brought forth when we remember. “

St. Augustin d’Hippone, Book X, Chapter IX, 397AD, Translated by J. G. Pilkington

~ ~ ~

« Là, ne s’arrête pas l’immense capacité de ma mémoire. Elle porte en ses flancs tout ce que j’ai retenu de la science, et que l’oubli ne m’a pas encore dérobé. Et ces perceptions, je les garde à l’écart plus intérieurement, non pas en lieu, ni en images, mais en réalité. Car ce que je sais de la grammaire et de la dialectique, du nombre et de l’espèce des questions, n’est pas entré dans ma mémoire comme l’image, qui laisse la réalité à la porte, évanouie aussitôt qu’apparue ; comme la voix imprimant à l’ouïe une trace qui la fait vibrer encore lorsqu’elle a cessé de raisonner ; comme l’odeur qui, dans son passage, dissipée au vent, pénètre l’odorat et porte à mémoire d’une image qui se reproduit au désir de la réminiscence ; comme l’aliment qui n’a plus de saveur qu’au palais de la mémoire ; ou comme l’objet que la main a touché, dont l’éloignement n’efface pas l’empreinte : car les réalités de cet ordre ne sont pas présentées à la mémoire, mais leurs seules images, qui, saisies avec une étonnante rapidité, sont rangées dans des cellules merveilleuses, d’où elles sont tirées merveilleusement par la main du souvenir. »

St. Augustin d’Hippone, Livre X, Chapitre IX, 397 apr. J.-C.

~ ~ ~

Résumé

La mémoire épisodique concerne la ré-expérience d'évènements personnels passés ancrés dans leur contexte d'encodage. Ces souvenirs épisodiques ne sont pas figés : leur contenu est influencé par la modalité sensorielle de l'indice de rappel. Par exemple, les souvenirs évoqués par des odeurs sont connus pour être moins fréquents, plus surprenants, vivaces, émotionnels, et anciens que les souvenirs évoqués par des images ou des mots. Ces phénomènes sont communément expliqués par les liens anatomiques étroits et directs qui existent entre les structures primaires olfactives, mnésiques et émotionnelles. Pourtant, les odeurs n'ont que trop rarement été comparées à des indices qui possèdent eux aussi des liens privilégiés avec la mémoire, comme la musique et les visages, tant au niveau comportemental que fonctionnel. Cette thèse présente deux grands objectifs : 1) Identifier et caractériser les particularités de la mémoire épisodique attribuables à la modalité sensorielle de l'indice de rappel (Etudes 1 et 2) ; 2) Etudier la dynamique des réseaux neuronaux qui sous-tendent le rappel épisodique et plus particulièrement les interactions qui sont modulées différemment selon la modalité sensorielle de l'indice de rappel (Etude 3). Pour tester l'hypothèse que l'émotion serait un facteur essentiel dans la particularité des indices olfactifs à remémorer un souvenir, cette thèse a pour objectif secondaire d'évaluer l'effet différentiel de l'émotion de l'indice de rappel épisodique en fonction de sa modalité sensorielle. Pour répondre à nos objectifs et permettre l'étude de la mémoire épisodique dans des conditions les plus écologiques possibles, nous avons développé un protocole de réalité virtuelle non-immersive déclinable en plusieurs versions permettant l'encodage et le rappel d'épisodes complexes et multisensoriels vécus en laboratoire.

Par l'utilisation de stimuli neutres, la **première étude** a montré qu'il existe une influence de la modalité sensorielle de l'indice de rappel sur les performances de mémoire de reconnaissance et épisodique. Les visages ont été très bien reconnus et ont été de très bons indices de mémoire épisodique ; les odeurs ont été moins bien reconnues mais ont été de bons indices de mémoire épisodique ; les musiques bien que très bien reconnues n'ont pas été de bons indices de mémoire épisodique. Par l'utilisation de stimuli émotionnels, la **seconde étude** a confirmé les résultats précédents, et précisé les effets de l'émotion sur les performances de mémoire épisodique en montrant que la valence émotionnelle de l'indice de rappel favorise globalement toutes les étapes de mémoire. Les stimuli les plus plaisants et déplaisants, par comparaison aux plus neutres, sont associés à de meilleures performances de mémoire. De plus, l'efficacité prononcée des odeurs à évoquer un rappel épisodique a été associée à la motivation à rééchantillonner le stimulus. Cette étude a également souligné l'importance de la pertinence écologique des stimuli, la virtualisation des visages conduisant à la suppression de

leur supériorité comme indice de mémoire vis-à-vis des odeurs et des musiques. **La troisième étude**, encore en cours, confirme la force des odeurs, quand elles sont plaisantes, à remémorer les différentes dimensions d'un épisode. Des données préliminaires, suggèrent que les indices musicaux et olfactifs de la mémoire épisodique activent les réseaux de la mémoire autobiographique. **En conclusion**, nos études montrent un effet de la modalité sensorielle de l'indice de rappel sur le rappel épisodique et suggèrent que cet effet est associé à l'émotion portée par ces indices. Les odeurs semblent des indices de rappels particuliers, associés à une performance de reconnaissance moyenne, mais favorisant une recollection précise des souvenirs épisodiques. Cette recollection est portée par la motivation qu'elles engendrent. Les musiques, bien que très bien reconnues, ne favorisent pas le rappel correct des dimensions épisodiques associées. Enfin, les stimuli visuels semblent différer en fonction de la pertinence écologique de ces derniers, une mémoire plus efficace et complète étant associée à des stimuli plus pertinents écologiquement.

Abstract

Episodic memory concerns the re-experience of past personal events anchored in their encoding context. These episodic memories are not fixed: their content is influenced by the sensory modality of the recall cue. For example, memories evoked by smells are known to be less frequent, more surprising, vivid, emotional, and older than memories evoked by images or words. These phenomena are commonly explained by the close and direct anatomical links that exist between the primary olfactory, memory, and emotional brain structures. However, odors have rarely been compared to cues that also possess privileged links to memory, such as music and faces, both behaviorally and functionally. This thesis has two main objectives: 1) To identify and characterize the particularities of episodic memory attributable to the sensory modality of the recall cue (Studies 1 and 2); 2) To study the dynamics of the neural networks underlying episodic recall and more specifically the interactions that are modulated differently according to the sensory modality of the recall cue (Study 3). To test the hypothesis that emotion would be an essential factor in the particularity of olfactory cues to recall a memory, the secondary aim of this thesis is to evaluate the differential effect of emotion of the episodic recall cue as a function of its sensory modality. To meet our objectives and to allow for the study of episodic memory in the most ecological conditions possible, we have developed a non-immersive virtual reality protocol that can be declined in several versions allowing the encoding and recall of complex and multisensory episodes experienced in the laboratory.

By using neutral stimuli, **the first study** showed that the sensory modality of the recall cue influenced recognition and episodic memory performance. Faces were very well recognized and very good cues for episodic memory; smells were less well recognized, but were good cues for episodic memory; musical excerpts, although very well recognized were not good cues for episodic memory. By using emotional stimuli, **the second study** confirmed the previous results, and clarified the effects of emotion on episodic memory performance by showing that the emotional valence of the recall cue favors globally all memory stages. The most pleasant and unpleasant stimuli, compared to the most neutral ones, were associated with better memory performance. In addition, the pronounced effectiveness of odors in evoking episodic recall was associated with participants' individual motivation to resample the stimulus. This study also highlighted the importance of the ecological relevance of the stimuli, with the virtualization of faces leading to the suppression of their superiority as a memory cue in comparison to odors and music. **The third study**, still in progress, confirms the memory strength of odors, when they are pleasant, to recall the different dimensions of an episode. Preliminary data suggest that musical and olfactory cues in episodic memory activate autobiographical memory networks. **In conclusion**, our studies reveal an effect of the sensory

modality of the recall cue on episodic recall and suggest that this effect is associated with the emotion carried by these cues. Odors appear to be singular recall cues, associated with average recognition performance, but favoring accurate recollection of episodic memories. This recollection is driven by the motivation the odors have generated. Music, although very well recognized, leads to less correct recall of associated episodic dimensions. Finally, visual stimuli seem to differ according to their ecological relevance, with more efficient cueing and more complete memory being associated with more ecologically relevant stimuli.

REMERCIEMENTS

Comme il n'est pas facile de raconter, de résumer tout ce qui a été développé, acquis, et conclu lors de la préparation d'une thèse, il est encore moins facile de remercier tous ceux qui ont contribué au développement et au bon déroulement de celle-ci. Ceux qui me connaissent savent que ma thèse ne s'est pas déroulée dans les conditions optimales. Elle a été difficile, tant au niveau des expériences qu'au niveau de la santé, qu'elle soit due au covid, psychologique, ou pire encore. Au milieu de tout ça, ma passion pour mon sujet de thèse m'a portée, mais la passion seule ne peut pas tout faire. Beaucoup de personnes ont été là pour moi, pour ma thèse, pour les expériences, pour réfléchir, pour apprendre et comprendre ensemble.

*Parmi ces personnes, mes directrices de thèse. Mes incroyables, mes géniales directrices de thèse, **Jane** et **Barbara**. Vous êtes des forces de la nature et avez été de vrais piliers à ce travail. Jane, nous interagissons depuis ma timide apparition dans ton bureau, lors de l'entretien pour un stage de Master 2. J'ai la scène bien en mémoire (épisodique !), une mémoire heureuse car j'allais être embarquée dans une merveilleuse aventure avec toi qui durera 5 années. J'ai été chaque jour comblée d'avoir partagé ce bout de chemin avec toi, et j'en sort plus grandie - scientifiquement et humainement - que je ne l'aurais jamais imaginé. Tu as été une alliée, supportrice, mentor, formatrice, guide, collaboratrice tant incroyable que je ne saurais trouver les bons mots, les mots justes comme tu les choisis si bien, pour décrire à quel point j'ai tant appris et tant aimé travailler avec toi. Ma reconnaissance est sans bornes, mille merci à toi. Barbara, quand tu as gentiment accepté de co-diriger cette thèse, j'étais aux anges ! Quelle chance d'interagir et collaborer avec une spécialiste de la mémoire musicale aussi gentille, disponible, que toi. Merci pour nos discussions scientifiques, tes remarques toujours très pertinentes, ton optimisme et ton soutien sans failles. J'ai été comblée, pendant cette thèse. Merci encore.*

*Je tiens également particulièrement à remercier l'association **ROUDNITSKA** pour son soutien financier et la confiance qu'il m'ont portée tout au long de mes quatre années de thèse. Merci !*

*Cette thèse est le fruit d'un travail collaboratif, et multidisciplinaire. Ainsi, je remercie de tout cœur mes équipes de recherche, **CMO** et **CAP**, qui m'ont permis de toujours apprendre plus, échanger et discuter, ouvrir mon champ de connaissance scientifique sur l'olfaction et l'audition. Je remercie également **Marc Thévenet**, qui a développé l'olfactomètre, et toutes les versions d'EpisOdor nécessaires au bon déroulement de nos expériences. Mille merci pour ta disponibilité à toute épreuve, ton professionnalisme, et ta bonne humeur malgré mes innombrables bugs de manips et oublis de dernière minute. Merci également à **Clément Désoche**, **Alessandro Farnè** et la plateforme **Neuro-Immersion** pour avoir développé le logiciel de réalité virtuelle Multisens, selon nos critères de plus en plus précis, et malgré mon perfectionnisme à toute épreuves. **Samuel Garcia**, mon cher Cassos en chef qui aime bien nous taquiner et râler avec sourire, je te remercie évidemment pour tes formations python éclair, pour ta disponibilité (surtout s'il y a du chocolat, haha !), pour m'avoir aidé à analyser mes données, réparer mes bêtises dans mes scripts et rattraper les bugs de manip, et pour avoir admis les*

forces de R par rapport à Python (un gros effort). Sans toi, je n'aurais jamais su coder. Et ça, je t'en suis infiniment reconnaissante. **Gaëlle Leroux**, tu as été une personne clé dans le design et l'analyse de l'expérience en IRMf. Les résultats sont très prometteurs, et c'est grâce à toi ! Merci aussi pour ta patience et tes explications claires et répétées sur tout ce qui touche à l'IRMf et nos analyses. **Phillipe Litaudon, Christelle Daudé**, et mes supers stagiaires **Eda Erdal, Théo Sommacal, Manon Lacour, et Mathilde Auvray**, mille mercis pour votre aide lors des expériences et de tous les pilotes qui les ont précédés ! Merci à **Brigitte, Florence, Hervé, Pierem, Romain** pour avoir tant facilité le déroulement de ma thèse, tant au niveau informatique qu'administratif.

Chers membres de mon comité de suivi de thèse, **Christopher Moulin** et **Gaën Plancher**, je vous remercie pour votre soutien et vos remarques avisées lors de nos entrevues annuelles.

Je remercie également les membres de mon jury, **Sylvain Delplanque, Mathilde Groussard, Christopher Moulin, Fabien Perrin**, et **Pascale Piolino**, pour leur présence, leurs questions et remarques toutes plus intéressantes les unes que les autres. J'ai été fière de défendre ma thèse devant vous, et très touchée par vos remarques positives et encourageantes sur mon travail.

Merci à Mme Labaste, ma professeure du lycée Gauguin qui m'a donné goût à la SVT, ainsi qu'à Remi Gervais, dont l'immense passion pour les neurosciences m'a contaminée. Merci à Lise Boisselier, qui m'a gentiment acceptée à ses côtés (même à son bureau !) et qui m'a initiée à la recherche, lors de mon stage de Master 1.

Merci à tous mes copains, mes amis, mes précieux soutiens. Merci aux étudiants de Neuropop, les vistavinatiens, et de CMO pour avoir rendu mon quotidien infiniment plus fun... et pour le film de thèse ! Merci à mes ptis biscuits pour tout ce que vous avez apporté, et qui ne se résume évidemment pas en deux lignes, vous le savez. Merci aux anciens du DUT GTE, les poulets coco. Merci à la team de Tahiti, Coline, Constance, Elodie, Hoani et Margaux, depuis le lycée, vous êtes là, et ce n'est pas rien de me supporter depuis tout ce temps. Merci à mes compagnes d'aventures, ma super cousine Eline, et Manon, pour m'avoir fait voir le monde et oublier ma thèse le temps de merveilleux voyages. Merci à Matthieu pour ses cartes de voyage et son soutien de loin. Un jour, j'irai te voir à Mayotte, promis !

Merci à ma famille, proche et plus lointaine, avec les étoiles. Papa, maman... sans votre curiosité scientifique et du monde, sans les nombreux voyages que vous m'avez fait vivre, sans nos deux « expatriations » à Tahiti, sans votre générosité et votre soutien, je n'aurais jamais eu l'occasion d'écrire ces mots.

Merci, chers sujets, qui m'avez fait confiance.

Merci à tous.

Scientific communications

COMMUNICATIONS ORALES

Rey, L., Désocche, C., Thévenet, M., Garcia S., Tillmann, B., Plailly J. (2018). Influence de la modalité sensorielle de l'indice de rappel et de son émotion sur le rappel épisodique. *Groupement Odorant-Odeur-Olfaction 2019, Toulouse, France*

POSTERS

Rey, L., Désocche, C., Thévenet, M., Garcia S., Tillmann, B., Plailly J. (2019). Are odors, musical pieces and faces equal to evoke episodic memories? A new encoding approach using virtual reality. *Groupement Odorant-Odeur-Olfaction 2019, Toulouse, France*

Rey, L., Désocche, C., Thévenet, M., Garcia S., Tillmann, B., Plailly J. (2018). Are odors, musical pieces and faces equal to evoke episodic memories? A new encoding approach using virtual reality. *Dynamic and flexible nature of memory 2019, Lyon, France*

PUBLICATIONS

Rey, L., Desocche, C., Thévenet, M., Garcia S., Tillmann, B., Plailly J. (*to be submitted*). Different patterns of recognition and episodic memory for odors, music, and faces – an ecological laboratory approach using virtual reality. *Memory*.

Rey, L., Desocche, C., Thévenet, M., Garcia S., Tillmann, B., Plailly J. (*in prep*). The effect of emotion and sensory modality on recognition and episodic memory for odors, music and faces.

Rey, L., Tillmann, B., Plailly J. (*submitted*). Characterizing emotional response to olfactory, auditory, and visual stimulations in a virtual reality environment. In: Bensafi (Ed.) *Characterization of emotions induced by smells, tastes and food. Methods and Protocols in Food Science*.

Table of content

ABSTRACT	10
SCIENTIFIC COMMUNICATIONS	12
TABLE OF CONTENT	15
TABLE OF FIGURES	17
ABBREVIATIONS	18
FOREWORD	19
SCIENTIFIC CONTEXT	21
1 MEMORY	23
1.1 A GLOBAL DEFINITION	24
1.2 HISTORICAL PERSPECTIVE ON MEMORY	25
1.2.1 Memory as a unitary process	25
1.2.2 Memory as a modular process	27
1.3 THE CURRENTLY ACCEPTED MODELS OF MEMORY	29
1.3.1 Short-term memory	29
1.3.2 Long-term memory	30
2 MEMORY OF PERSONAL EVENTS	33
2.1 MAIN THEORIES	34
2.1.1 The Episodic framework	34
2.1.2 The Autobiographical framework	37
2.1.3 Present definition of episodic memory	39
2.2 EXPERIMENTAL APPROACHES	40
2.2.1 Laboratory approach: Episodic memories studies	40
2.2.2 Ecological approach: Autobiographic memories studies	43
2.2.3 Laboratory-ecological episodic memory approaches	46
2.3 NEURAL BASES OF MEMORY OF PERSONAL EVENT	48
2.3.1 Medial temporal lobe	48
2.3.2 Cortical pathways	50
2.3.3 Influence of the experimental approach on neural bases of personal events	52
3 MEMORY AND THE SENSES - HOW SENSORY MODALITIES INFLUENCE MEMORY OF PERSONAL EVENTS	53
3.1 THE CASE OF FACE	54
3.1.1 Importance of faces in human everyday life	54
3.1.2 Face perception, emotion and cognition	55

3.1.3	Face and memory	60
3.2	THE CASE OF MUSIC	63
3.2.1	Importance of music in human everyday life	63
3.2.2	Music perception, emotion, and cognition	64
3.2.3	Music and memory	69
3.3	THE CASE OF ODOR	74
3.3.1	Importance of odors in human everyday life	74
3.3.2	Odor perception, emotion and cognition	74
3.3.3	Odor and memory	79
OBJECTIVES & HYPOTHESIS		85
RESULTS		91
STUDY 1		93
EPISODIC MEMORY FOR ODORS, MUSIC AND FACES— AN ECOLOGICAL LABORATORY APPROACH USING VIRTUAL REALITY		93
STUDY 2		147
EPISODIC MEMORY EVOKED BY ODORS COMPARED TO MUSICAL EXCERPTS AND FACES: MODALITY-SPECIFIC EFFECT OF WANTING, BUT MODALITY-UNSPECIFIC EFFECT OF VALENCE		147
STUDY 3: NEURAL BASIS OF EPISODIC MEMORY FOR EMOTIONAL ODORS AND MUSIC		189
GENERAL DISCUSSION		211
1	A NEW PROTOCOL TO INVESTIGATE EPISODIC MEMORY	213
2.1	ODORS	216
2.2	MUSIC	218
2.3	FACES	219
2.4	POTENTIAL EXPLANATION FOR DIFFERENCES OBSERVED BETWEEN ODOR, MUSIC AND FACE CUES	220
3	INFLUENCE OF SENSORY EVALUATIONS ON MEMORY PERFORMANCE AS A FUNCTION OF SENSORY MODALITY	221
3.1	IS THE EMOTION EFFECT THE SAME FOR EACH SENSORY MODALITY?	222
3.2	IS SENSORY MODALITY INFLUENCED BY OTHER SENSORY DIMENSIONS ?	224
CONCLUSION		225
REFERENCES		231

Table of figures

<i>Figure 1: Mnemosyne – A figure of memory in ancient Greek mythology</i>	23
<i>Figure 2: Memory as a global process</i>	25
<i>Figure 3: Systemic division of human long-term memory according to the Squire Model</i>	32
<i>Figure 4: The recall of past events memories is a reconstructive process</i>	33
<i>Figure 5: The MNESIS model : an integrative model of memory</i>	36
<i>Figure 6: The self-memory system</i>	38
<i>Figure 7: Schema of the Binding of Items in Context (BIC) model</i>	49
<i>Figure 8: Posterior Medial – Anterior Temporal (PMAT) framework</i>	51
<i>Figure 9: The Sensorama Simulator (Heilig, 1962)</i>	53
<i>Figure 10: Unfamiliar face perception</i>	55
<i>Figure 11: The visual pathway, from the eyes to neocortical areas</i>	58
<i>Figure 12: Face-selective areas in a typical participant</i>	59
<i>Figure 13: The auditory pathway, from the auditory nerve to neocortical areas</i>	67
<i>Figure 14: Cortical areas dedicated to the musical processing of information</i>	68
<i>Figure 15: Cortical areas dedicated to the emotional processing of musical information</i>	69
<i>Figure 16: Schematic view of the human olfactory system.</i>	77
<i>Figure 17: AM distribution across lifespan as a function of sensory modality</i>	81
<i>Figure 18: Sensory modality effect on the number of defining words</i>	221

Abbreviations

BRAIN AREAS

-  **AG:** angular gyrus
-  **AON:** anterior olfactory nucleus
-  **Hipp:** hippocampus
-  **(L/M)EC:** (lateral/medial) entorhinal cortex
-  **IPFC:** lateral prefrontal cortex
-  **mPFC:** medial prefrontal cortex
-  **NAcc:** accubems nucleus
-  **OB:** olfactory bulb
-  **OFC:** orbitofrontal cortex
-  **OT:** olfactory tubercle
-  **PC:** piriform cortex
-  **PCC:** posterior cingular cortex
-  **PHC:** parahippocampal cortex
-  **PRC:** perirhinal cortex
-  **RSC:** retrosplenial cortex
-  **STS:** superior temporal sulcus

OTHER ABBREVIATIONS

-  **AM:** autobiographical memory
-  **CR:** correct rejection
-  **EM:** episodic memory
-  **FA:** false alarm
-  **LTM:** long-term memory
-  **MMS:** multiple memory system (theory)
-  **R/K:** remember/know (paradigm)
-  **SDT:** signal detection theory
-  **SM:** semantic memory
-  **STM:** short-term memory
-  **VR:** virtual reality
-  **WWW:** What-Where-When/Which Context (paradigm)

Foreword

I've spent a part of my life in Tahiti as a child and then as a teenager. I really loved to live there, to be embedded in this culture, and I always remember these times with nostalgia and a little of sadness. At some points in my life, I would love to forget this missing part of me but in the end, I am always grateful to still have these memories. These events have shaped my life, the way I perceive the world. Without the memory of these earlier times, I would not have been who I am today but a fundamentally different person.

I recently bought a little bottle of *Monoï* made with *Tiare* flower from one specific brand coming from Tahiti. At this point, I did not realize that each time I smelled it I would feel like I was transported in Tahiti, even in the middle the entirely different atmosphere of Lyon. When this happens, I would love to have the power to keep this smell with me and hold onto those reassuring memories for much longer. This odor is not "just an odor", it's a place, a feeling, and a story. The intensity of this feeling never reached such height with any other kind of stimulus, even an olfactory stimulus. Thus, I wondered, is this phenomenon specific to this odor? Does it come from the link I created with it or is it due to the olfactory system? A similar feeling can be found with Tahitian music. It also moves me when I hear it but not (quite) as much. Why is this different than the one with the *Monoï* odor?

This thesis will answer some of these questions. In our work, memories of past events evoked by odors will be compared to those evoked by music or faces. Comparing memories evoked by different senses require to define memory and what distinguishes the memory of our past events from other forms of memory, to explain how different sensory modalities are perceived, memorized, and used as a cue, and what are the existing methods that can be used to compare them. To this end, a first chapter (1. Memory) will present how the definition of memory has evolved throughout ages to the current models of memory. The second chapter (2. Memory of personal events) will explain the two frameworks and related experiments on which the memory of past events is defined, and how these frameworks can be gathered into experiments that mimics reality. The third chapter (3. Memory and the senses) will provide an overview of perception, memory and neural bases particularities of face, music and odor stimuli. I will then present the questions and hypothesis that have been motivated by all the works presented across these three chapters (Objectives & Hypothesis). In the following parts I will present the three studies that have been performed to test these hypotheses (Results) before discussing the whole doctoral work and concluding (General Discussion).

SCIENTIFIC CONTEXT

1 Memory



Figure 1: Mnemosyne – A figure of memory in ancient Greek mythology

Mnemosyne is a titaness, the goddess of memory in the Greek mythology. It is the daughter of Ouranos (the sky) and Gaia (the earth) and the mother of the nine muses. The word Mnemosyne comes from the Greek word *mnēmē*, which means "remembrance, memory". (Rights: *Mnemosyne, de moeder der Muzen, Johann Theodor de Bry/CC0 1.0*)

As personified by the figure of Mnemosyne, the Greek goddess of memory, and as observed in everyday life, memory seems to be unique. There seems to be one memory. In fact, this perception of memory is old and has shaped memory research and history from Antiquity until the 20th century, when its definition has changed the most. To understand how the concept of memory evolved and how memory is now defined, this chapter will be divided into three sections. In the first section (1.1) a global, definition of memory will be provided. The next section (1.2) will present how the concept of memory has evolved throughout ages, from its perception as a single phenomenon (1.2.1) to a complex, varied and rich concept (1.2.2). Finally, the last section (1.3) will provide the current widely presented model of memory.

1.1 A global definition

Almost all memory researchers and computer scientists agree to the definition of memory as a serial process allowing the encoding, storage and finally – maybe – retrieval of information (Ebbinghaus, 1885; Melton, 1963). In human memory research, the **encoding** of information refers to the selection and the transformation of a labile, transient, sensory-perceptual trace coming from the environment (or the brain) onto a code, an engram. Encoding can be operated implicitly when it occurs unconsciously or explicitly otherwise. The involvement of **consolidation** processes will allow for its storage by transforming the labile sensory-perceptual trace into a more stable one. These consolidation processes are separated into two levels: synaptic consolidation occurs with molecular, cellular and synaptic modifications, and systems consolidation occurs with neural networks rearrangements (for reviews, see: Dudai, 2004; Dudai *et al.*, 2015). The **storage** of the encoded information can last from few seconds to months, decades, or even the entire life. Then, sometimes, consciously or not, the information is retrieved. The **retrieval** implies the restauration of the memory trace in its first labile form and is thus sensitive to the retrieval environment (mood or emotional state, perceptual or semantic information). For example, cue emotion can change both the access and content of our memories (Sheldon and Donahue, 2017; Sheldon *et al.*, 2020). The retrieved information is then forgotten or goes through another consolidation, process called **re-consolidation**, to be stored again (Nader *et al.*, 2000). When re-consolidation occurs, the restored memory, but also all the information that had been encoded while memory was retrieved are susceptible to be consolidated and stored into the same memory (St. Jacques and Schacter, 2013).

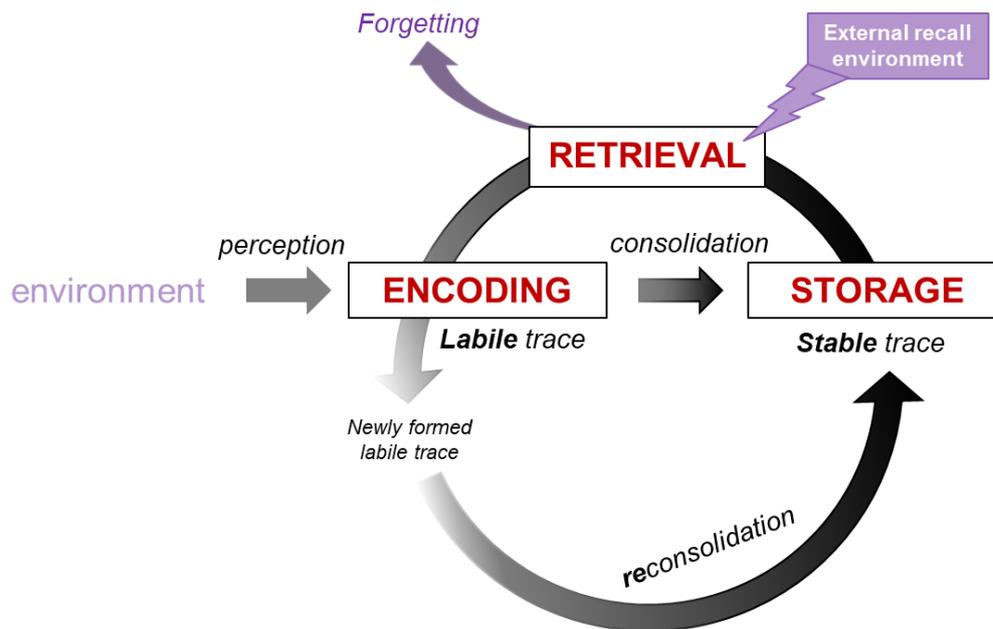


Figure 2: Memory as a global process

All forms of memories follow the same general steps. Information coming from the environment is first perceived and encoded. This labile trace needs then to be consolidated to be stored. The recall of this information implies restoration into their first, labile form. This change of state makes this information sensitive to forgetting and to external cues. After the recall the information, that may have been changed is then reconsolidated to be stored again.

1.2 Historical perspective on memory¹

The global definition of memory presented previously explains the processes that allow information to be integrated into memory. However, this definition is very large and gathered many sub-processes, many views of what a memory is. Indeed, the definition of memory has evolved from antiquity until now thanks to the contribution of many disciplines, among them: philosophy, psychology, neuropsychology, neuroscience, and cognitive science.

1.2.1 Memory as a unitary process

Memory has been the subject of human curiosity over centuries, and its definition evolved throughout the ages. During Antiquity, memory appears as a goddess called Mnemosyne (Figure 1). At this time memory was seen as a unitary and rigid process, allowing for an exact reproduction of past events. For example, Plato compared memory to an impression on a wax tablet. Memory was understood as the way allowing people to gather knowledge, recite a poem, or a theater text. Over the following centuries, numerous scholars tended to define memory, dividing it in innate *versus* acquired memory, or sensible *versus* intellectual memory (St Augustin d'Hippone, 397AD), body (inconscient) *versus* intellectual memory (Descartes, 1640, see: Fóti, 2000). This early work was of importance in contemporary theories of memory,

¹ This section had been written with the help of the exhaustive book of Francis Eustache and Béatrice Desgranges (2010)

notably providing the first basis to construct experiments, understand memory pathologies, and developing models of memory.

Many modern concepts and paradigms had been set up from 1880 to 1915, which can be seen as the first golden age of memory research. Research on memory focused first on forgetting and memory loss, as evidenced by the paradigms that had been set up. The most well-known work of this period is the very first laboratory studies on memory by Hermann Ebbinghaus (1885; translated version: Ebbinghaus *et al.*, 1913). Thanks to numerous experiments testing himself using non-words, he developed the *saving method* which measures the time necessary to relearn forgotten material and describes the occurrence of a forgetting rate. Hermann Ebbinghaus also underlined the existence of a non conscient memory and was the first to introduce the three major phases of all memory processes: “learning, retention and recitation/reproduction”. These phases were until now re-used in the framework of information processing theory of cognition, including memory, in experimental works (*e.g.*, Chen *et al.*, 2015; Melton, 1963) and textbooks (*e.g.*, Gazzaniga *et al.*, 2019). Other important advances in experimental memory research were set up during this period. For example, Mary Whiton Calkins developed in 1894 the paired associate paradigm, which enables to study the effect of sensory modality (vision and audition), recency, and vividness on short term memory of association (see Madigan and O’Hara, 1992 for an overview of her works). In the 1900’s, Edmund Husserl (Husserl and Gibson, 1931) created a new science he referred to as phenomenology, motivated by the need to study conscious and subjective phenomenon in a neutral and objective way. Phenomenology has been widely used in memory studies, especially to characterize the experience of remembering memories and not just the content of the memory. For example, remembering a past event can be experienced with the impression of reliving the event, or just by knowing what happens.

The study of amnesic patients enabled considerable theoretical advances in memory research. In 1881, Théodule Ribot, a French philosopher and physiologist, demonstrated that memory is evolving with time by the study of amnesic patients. He stated the Ribot’s Law that describes memories as being first instable, weak and unorganized, and then becoming more and more stable, automatic and highly organized with time. For example, amnesic patients’ memories’ destruction begins with recent and more unstable past memories to finish with more stable ones. The investigation of amnesic patients suffering for the Korsakoff syndrome led Sergeï Korsakoff (1889) to confirm Ribot’s observations and a distinction between ancient and recent memory. Korsakoff noted that his patients seemed to be always living at the same moment, the onset of the disease, but kept a kind of sense of familiarity, he indeed noted a difference between a memory of feelings and unconscious impressions, and a memory of time, locations and forms. Edouard Claparède, a Genevan neurologist, later studied memory of these patients

(1907). Notably, he re-used the saving method set up by Hermann Ebbinghaus and showed that even though Korsakoff patients were unable to consciously remember learning items, their response time decreased after relearning.

Based on these works, philosophers and researchers began to propose that memory may be divided onto different forms. Theodule Ribot (1870) explained that memory of sensations is different than what he called a memory of ideas. Edouard Claparède provided the first experimental evidence of the ability to learn without being aware of it, bringing support to Sergeï Korsakoff (1889) hypothesis of the existence of an unconscious memory. In his book *The principles of psychology*, William James (1890) separated memory into different stores based on the retention duration: a short-term store (primary memory), and a long-term store (secondary memory). Although his work was purely theoretical, the distinction between short-term and long-term memory remains valid today. As we will see below, famous neuropsychological and lesion studies, altogether with new investigation methods, have confirmed observations of different forms of memory.

1.2.2 Memory as a modular process

In 1953, Henri G. Molaison, also known as H.M. patient, underwent a bilateral surgery of the bilateral medial temporal lobes (**MTL**) to be cured of a severe epilepsy. After the surgery, he had lost many mnemonic abilities and showed an anterograde and a partial retrograde amnesia, while other faculties remained preserved. He had no deficit in general intelligence, perception, abstract thinking and reasoning ability. His case was studied by William Scoville and Brenda Milner, and their research on H.M and other patients resulted in a first article published in 1957, showing the importance of the hippocampus (**Hipp**) and the para-hippocampal gyrus in the memory of recent events (Scoville and Milner, 1957). Brenda's Milner work on patient H.M opened the gate on a new era of research for which memory was not a unitary process. More and more researchers aimed to understand, with neuropsychological and lesion studies, how information from the environment is stored in the brain, and to investigate the neural bases of memory. The base idea of their work can be summarized by the concept of (simple) **dissociation**, relying on the hypothesis that each form of memory is underpinned by a cerebral area. Thus, if a memory area is lesioned, then the behavior supported by this region can no longer be expressed (Eustache and Desgranges, 2010; Ferbinteanu, 2019).

The first influential cognitive model explaining memory as a process was the *modal model*, presented by Richard Atkinson and Richard Shiffrin (1968; for an example of its influence, see: Baddeley *et al.*, 2019). Influenced by the unitary view, this model explains memory as being essentially serial. Information from the environment can be stored in three components guided

by control processes: a sensory register, a short-term store, and a long-term store. The large amount of information gathered from the environment is first selected by the sensory register, which can store information for less than one second. Selected information is then sent to the short-term store, which can keep it for about 30s. If the information is *actively* maintained, it enters the long-term store and can be stored the entire life. To be rehearsed, information must be processed again by the short-term store, which selects the needed information through control processes. This model has been under criticisms, one of them being the passage from the short-term to the long-term store. The *modal model* suggests it to rely mainly on conscious rehearsal (Camina and Güell, 2017; Plancher and Barrouillet, 2020), but there are other explanations that describe retention in the long-term store. For example, Fergus Craik (Craik, 2020; Craik and Lockhart, 1972; Craik and Tulving, 1975) explains it with the *level of processing* theory, the fact that adding meaning to an information implies a better processing, and recall.

The ability to store information in long-term store without a necessary step in a short-term store was demonstrated later by the study of patient K.F (Shallice and Warrington, 1970). Patient KF had an impaired short-term memory (**STM**) while having an intact long-term memory (**LTM**) after a brain trauma. The description of his memory impairment showed a **double dissociation**, that is the demonstration of an opposition between two cognitive behaviors and two lesioned brain regions. Thus, HM showed an intact STM and an impaired LTM, while K.F showed an intact LTM and an impaired STM. These works on simple and double dissociations on neuropsychological cases, but also on animal studies (Squire, 1992; or see the review of Ferbinteanu, 2019), highly contributed to the hypothesis of a systemic organization of memory and provide the first demonstration of two systems that are distinguished on the basis of retention duration, as proposed earlier by William James (1890).

Following the demonstration of these dissociations, a currently widely reported theory of a systemic memory has been proposed by Larry Squire and Stuart Zola-Morgan (Squire and Zola-Morgan, 1988), named the **multiple memory system (MMS)**; for reviews, see Squire, 2004; Squire and Wixted, 2011). Basically, with the knowledge gathered through the 20th century, they proposed that memory can be divided into a memory that is expressed through language (declarative memory) versus a memory that is expressed through performance (non-declarative memory). They further proposed the **medial temporal lobe memory system theory** (Squire and Zola-Morgan, 1991) that explains the crucial role of medial temporal lobe in establishing LTM for facts and events. These two theories have evolved since their first publication (see for review: Ferbinteanu, 2019), but they are still the basis of current theory and are still taught in university classes (see for example: Gazzaniga *et al.*, 2019).

1.3 The currently accepted models of memory²

Today, almost all memory models agreed on the fact that information coming from the environment is first processed in a sensory memory system before being processed into STM and/or LTM. **Sensory memory** refers to the storage of raw (or weakly processed) data coming from the environment, that are maintained just long enough to allow most relevant information to be selected, mainly by attention processes (Atkinson and Shiffrin, 1968). Sensory memory has been well described for visual stimuli (iconic memory), less for auditory and tactile stimuli (respectively, echoic and haptic memories (Camina and Güell, 2017)), and not for olfactory stimuli.

1.3.1 Short-term memory

Short term memory theory has been constructed around verbal, visual and auditory sensory modalities (Annett, 1996). The name given to this system refers to the duration of its storage: the information is kept for a short period of time. During this retention time, the information could be forgotten, processed, or (re)consolidated. Working memory is a sub-component of STM able to store information while it is manipulated to produce a complex behavior. It has been first theorized by Baddeley and Hitch (1974). They proposed a model called *the multi-component model of working memory*, which has been revised in 2000 (Baddeley, 2000; Baddeley *et al.*, 2019). Alan Baddeley and Graham Hitch suggest a four components system involving: a phonological loop, a visuospatial sketchpad, an episodic buffer, and a central executive. The phonological loop and the visuospatial sketchpad are two slave systems, as they behave like storehouse. Phonological loop is dedicated to the short-term storage and manipulation of verbal and acoustic information. All information that can be verbalized is here transformed into a phonological code. This loop also includes rehearsal processes, which refresh information to maintain it longer. The visuospatial sketchpad concerns visuospatial and kinetic information. These two slave systems are controlled by the central executive, which is a separate attentional control system. The central control system manages also STM interaction with LTM. The episodic buffer has been added lately to take into account the binding of information coming from other systems (including LTM). It is thus an associative component able to replace information in a spatio-temporal context, and to make information available to consciousness (Baddeley, 2000; Baddeley *et al.*, 2019; Camina and Güell, 2017). While being a general theory, this definition does not include material related to chemical senses, while an olfactory working memory does exist (Andrade and Donaldson, 2007; Moss *et al.*, 2019), nor does it include the specific characteristics related to musical material (Berz, 1995; Williamson *et al.*, 2010).

²This section was written with the help of the review written by Eduardo Camina and Francisco Güell (2017)

1.3.2 Long-term memory

LTM concerns information that can be stored for a very long period of time, from hours to life. The most famous systemic model of LTM, the MMS, proposes the existence of six different systems (Squire, 2004; Squire and Zola-Morgan, 1988). These systems have been firstly divided into two sub-categories, based on the recall measures and the reliance on MTL (or not): memory is either **non-declarative** when the outcome measure is a task performance or a change of behavior and does not engage MTL, or **declarative** when the memories are rehearsed with language and engage MTL. This distinction between non-declarative and declarative memory is also presented as an implicit (unconscious)/explicit (conscious) memory distinction, the non-declarative memory being understood as implicit and the declarative memory as explicit. This explicit/implicit distinction does not seem to be well suited, because for example some non-declarative memory may be rehearsed with consciousness (Cabeza and Moscovitch, 2013; Ferbinteanu, 2019; Henke, 2010; Madan, 2020). Following this first non-declarative/declarative division, the systems were classified according to the behavior they result in and their neural bases: Non-declarative memory has been divided into four systems, and declarative memory into two (Camina and Güell, 2017; Squire and Zola-Morgan, 1988)

1.3.2.1 Non-declarative memory

Non-declarative memory systems are grouped onto two bases: (1) their rehearsal is measured in terms of nonverbal performance, such as decreased reaction time, (2) the resection of MTL does not widely disrupt their rehearsal (Squire, 2004; Squire and Zola-Morgan, 1991, 1988). These memories had been defined as being *anoetic* by Endel Tulving (1985), that is, they do not require the conscious remembering of a past experience. Four non-declarative systems had been defined by Larry Squire and colleagues (Squire and Zola-Morgan, 1988):

- The **procedural memory system** involves a long and repetition-based learning. The resulting behavior is the acquisition of habits or skills, such as knowing how to drive a car, to write, or to always put the keys in their right place without paying attention.
- **Priming** is involved when the exposure to a stimulus influences later behavior. For example, after having read the word “doctor”, subsequent processing of the word “nurse” will be faster than processing of the word “biscuit”.
- **Simple classical conditioning** concerns two kinds of conditioning. *Operant (skinner) conditioning* involves the reinforcement or punishment of a behavior (giving food after a rat has clicked on a button will reinforce this behavior); and *pavlovian (classical) conditioning* in which a previously neutral stimulus (a bell) will acquire a reward or punitive value after repeated associations with another rewarding or punitive stimulus

(food or electric shock) (the dog of Pavlov learned that the ring of a bell signaled incoming food, so the dog salivated when the bell was ringing, before the food arrival).

- **Non associative learning**, such as habituation or sensitization, concerns respectively the decrease or increase of a behavior following repeated exposure. For example, habituation explains that when exposed to an odor for a long period of time, odor will be perceived less and less strong. On the contrary, sensitization explains for example that smelling repeatedly an odor that is poorly perceived (*e.g.*, being not intense) can increase its perception.

1.3.2.2 Declarative memory

Declarative memory systems are regrouped onto two bases, that were historically defined by Endel Tulving (Tulving, 1972): (1) their rehearsal can be shared through language, (2) the resection of MTL does disrupt their rehearsal (Squire, 2004; Squire and Zola-Morgan, 1988).

Two declarative systems had been defined by Squire and colleagues:

- **Semantic memory (SM)** is the memory of facts, concepts and general knowledge. It is acquired through repetition like when a theatre piece is learned and remembered by an actor. These memories had been defined as being *noetic* by Endel Tulving (1985), their recall comes with the awareness of remembering known facts, but without any sense of having learned them. Semantic memory allows for communicating, manipulating even highly abstract concepts, structuring our memories, *etc.*
- **Episodic memory (EM)** is a system that is perceptually based, by opposition to SM. It represents the recollection of memories and has been defined as the memory of past and future events embedded in a spatio-temporal context. These memories had been defined as being *autonoetic* by Endel Tulving (1985). Their recall involves the subjective feeling of remembering personal memories that are oriented toward the past or the future.

However, this subdivision of declarative memory into two subsystems cannot explain all memory phenomenon. For example, memory of personal events, or autobiographical memory (AM) may contain both episodic and semantic information. Most of the recollected memories are made of factual and perceptual details, that are both combined to make memories rich and informative. As we will see in the next chapter, the memory of past event can be understood and studied with two points of view.

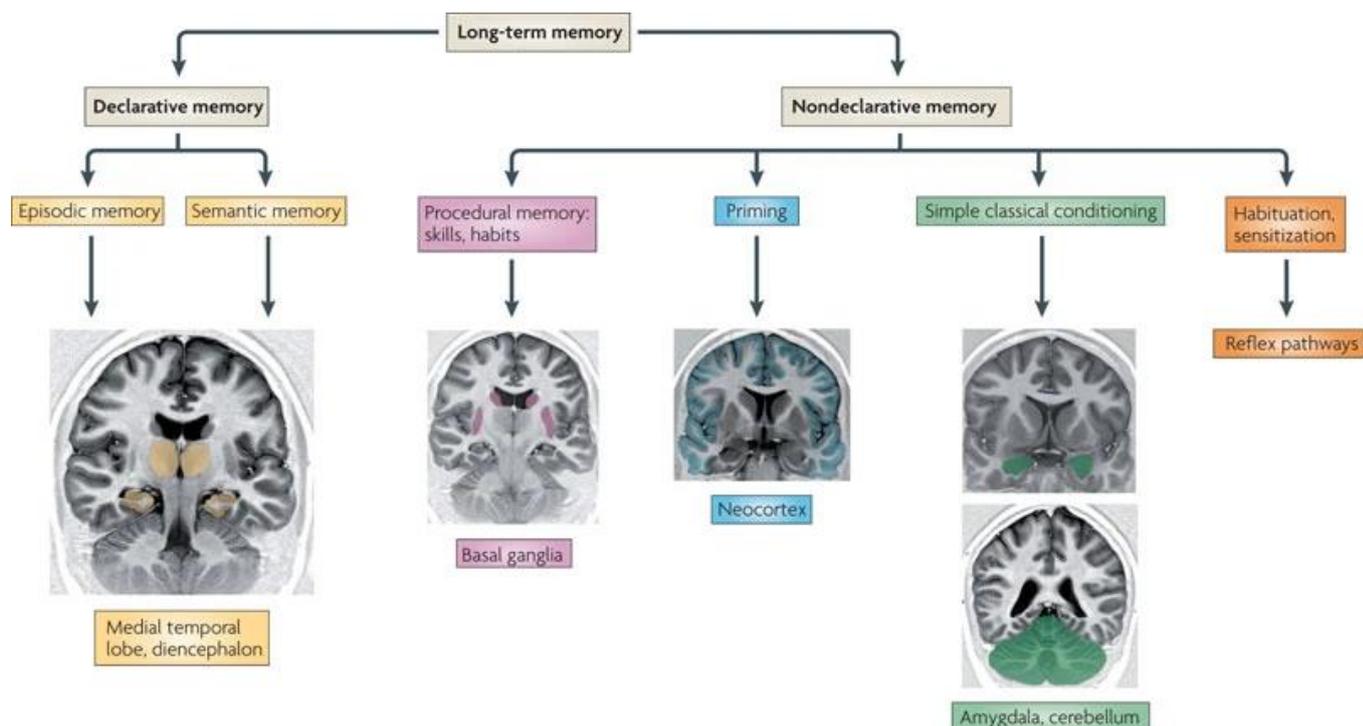


Figure 3: Systemic division of human long-term memory according to the Squire's Model

In this model, long term memory is divided into two categories. Declarative memory encompasses memory systems in which memory can be retrieved through language: semantic memory allows the retrieval of facts and concepts, and episodic memory is the memory of personal events. Nondeclarative memory encompasses memory systems in which memory cannot be retrieved through language but via motor action: procedural memory concerns skills and habits, priming concerns the stimulus influence on a later behavior, the conditioning concerns associative rewarding/punishing learning, and non-associative learning such as habituation and sensitization. *(Reproduced with permission via Copyright Clearance Center: Henke, 2010; Copyright © 2010 Nature Publishing group)*

2 Memory of personal events

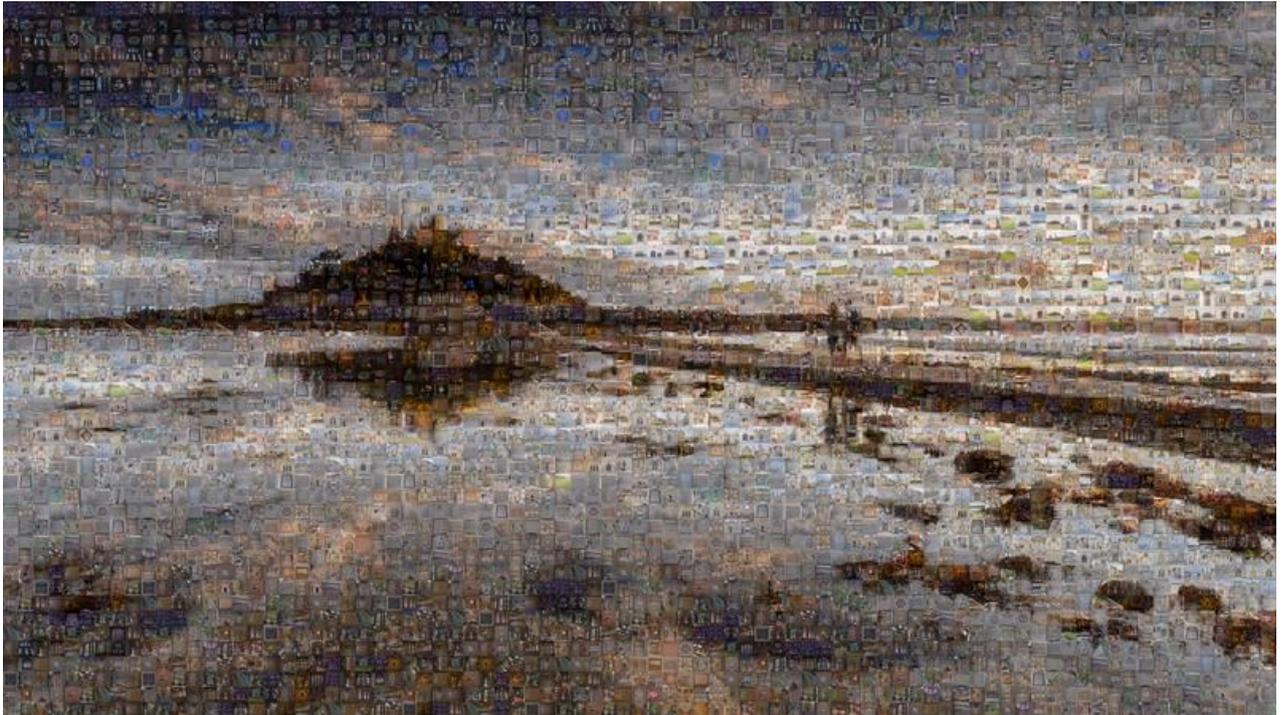


Figure 4: The recall of past events memories is a reconstructive process

This image is a Photomosaic of the St Michael's Mount. Photomosaic technic consists of creating a meta-image by the association of smaller images. With the eye of a memory scientist, the image can represent the reconstruction of a personal memory, as memories are reconstructed by reassembling previously encoded items into a coherent image.

(Rights: Fuzzypiggy," Photomosaic WLM2014 Winner". In: Wikimedia commons [online. Available from: https://commons.wikimedia.org/wiki/File:Photomosaic_WLM2014_Winner.png. CC BY-SA 3.0)

As presented in the image above (Figure 4), personal memories are complex, dynamic, and reconstructive. They are made of many and diverse representations that are combined to form coherent pictures that allow people to construct their own history and imagine their future according to their current situation. Two major memory systems contribute to their encoding, storage and retrieval: SM and EM. For example, I can recollect with details the memory I have of our second expatriation in Tahiti, in 2007, when my brother and I were so tired that we locked our parents out of our apartment before going to an involuntary, deep sleep. This memory contains episodic details such as our surprise when we woke up, but also semantic details such as "our second expatriation in Tahiti, in 2008". SM and EM are two separate systems in Squire global model of memory (see part 1.2.2; Squire, 2004; Squire and Zola-Morgan, 1991). They are thought to depend on different neural networks and mental processes, and are characterized by different rules and characteristics. This separation into systems does not exclude that both systems can work together to encode and/or retrieve memory of events, depending on task requirement. Some theorists have tempted to describe how EM and SM are associated to form autobiographical memories (Conway *et al.*, 2019; Eustache and Desgranges, 2010; Renoult and Rugg, 2020; Tulving, 1995), and two global frameworks

resulting from these works were set up. These frameworks are related to two outlooks of what AM is, that are represented by two different types of paradigms. This chapter will first present these two frameworks explaining how AM can be constructed with the interaction of EM and SM, before developing the definition of EM we will use in this thesis (2.1). Then we will see current experimental paradigms used to investigate AM and EM (2.2), and finally the neural bases that underpin AM and EM (2.3).

2.1 Main theories³

2.1.1 *The Episodic framework*

According to Endel Tulving, EM is the function allowing the encoding, storage, and retrieval of personal events that are recalled in their spatio-temporal contexts of encoding (Tulving, 1972). A cue will first produce an *ecphory*, a concept invented by Richard Semon in 1904 and re-used by Endel Tulving. *Ecphory* is a process through which a cue interacts with stored information, which will be elaborated and then reconstructed to be consciously available through **recollective experience**, the mental re-experience of the event with the conscious awareness to have lived it (Tulving, 1982; for a recent review, see: Staniloiu *et al.*, 2020). The notion of consciousness is of particular importance in Endel Tulving's theory, and is the key notion that distinguishes between EM and SM (Tulving, 1982, 1985). Let's take for example the following memory: "I was at school in Papeete, I was 15 so it was in 2007, I was bored because the course of mathematic was too easy, so I looked outside because that looks cool". I can *know* that this event happens (noetic consciousness), making this memory a SM, or truly *recollect* it with the subjective feeling of reliving it (EM, auto-noetic consciousness).

Phenomenological characteristics associated to EM are of great importance in Endel Tulving's works. As we have seen, they are the basis of the distinction between EM and SM, but it is also, according to Endel Tulving, a point that separates human and non-human animals (see for example: Tulving, 2002). Inspired by Endel Tulving's works, Nicola Clayton and Anthony Dickinson (1998) formalized EM definition by three core dimensions: the object of an event (*What* happened), its spatial location (*Where*), and its temporal location (*When*). This reduced definition allowed them to show that non-human animals were capable of remembering *What-Where-When* information as a whole after encoding it only one time (Clayton and Dickinson, 1998). However, the inability to prove animal auto-noetic consciousness while remembering, prevented them to conclude that animals recalled information through EM. Thus, animals are told to have an *episodic-like memory*. While being quite old, the debate is still standing, and

³ This part has been written with the help of the Handbook of Research Methods in Human Memory (2018)

some authors recently proposed to redefine EM without that consciousness part, allowing animals to have an EM (Madan, 2020).

Endel Tulving suggested that even if SM and EM belong to different systems, they are highly inter-related (Squire, 2004; Squire and Zola-Morgan, 1991). So, in 1995 he set up the Serial-Parallel-Independent (**SPI**) model, that he refined in 2003 in a personal communication (Eustache and Desgranges, 2010; Tulving, 1995). He first developed the hypothesis that procedural memory entails SM which entails EM, each associated with a more and more specialized state of consciousness: respectively, anoetic, noetic, and auto-noetic (Tulving, 1985). In the SPI model, Endel Tulving focused on what he called cognitive system, thus excluding the procedural system. The SPI model explains how cognitive systems are working together as a process, through encoding, storage and retrieval phases of memory (Tulving, 1995). The encoding of information is serial: information coming from the environment must first be unconsciously processed by a perceptual representation system, then by SM and finally by EM. The treated information is then stored and recalled in parallel, explaining cases of patients who have impaired SM but could still recall episodic events. The serial encoding of information from SM to EM is explained by the fact that information that is encoded deeper is associated to better performance (Level of processing theory, see Craik, 2020). However, this hypothesis is too restrictive. For example, it predicts that if SM is impaired, so does EM, but this is not the case (Graham *et al.*, 2000).

Taking those criticisms into account and including Baddeley's theory in a broader model, Béatrice Desgranges and Francis Eustache proposed the **Memory NEO-Structural Inter-Systemic model (MNESIS)**; Eustache and Desgranges, 2008, 2010, Figure 5). Their model provides a hypothesis to understand how perceptual memory, SM, EM, STM, and procedural memory interact to encode, store, and retrieve personal memories. It focuses on the dynamic and reconstructive nature of memories, including the ability of false memory integration and the semanticization of episodic memories, and can be read on different time scales (Eustache and Desgranges, 2010). According to MNESIS, information is encoded, consciously or not, similarly as in the SPI model. Most information coming from the environment is first processed in perceptual memory, then in the SM and the EM, while some information can bypass SM and be only processed by perceptual and EM. The information coming from these three systems can be used by working memory thanks to the episodic buffer, which acts as a link between LTM and STM. It can also be used by procedural memory, which can extract repeated patterns, perceptual-motor or cognitive procedure, from information coming from SM, EM and perceptual memory. Importantly, this model, by two retroactive rows, explains the recollection and semanticization processes. As recollection implies that stored information returns into its first labile form, MNESIS involves that EM stored information can be re-processed by perceptual

memory. A process represented by an arrow going from EM to SM explains the semanticization over time of information, that are for example conceptualized with time. This semanticization process can also occur with repeated retrieval as a recollection involves information to be re-encoded, so to be treated another time by SM. As most information are processed by SM before being processed by EM, MNESIS model explains why most recollections contain semantic information and are based on concepts or representations constructed via SM.

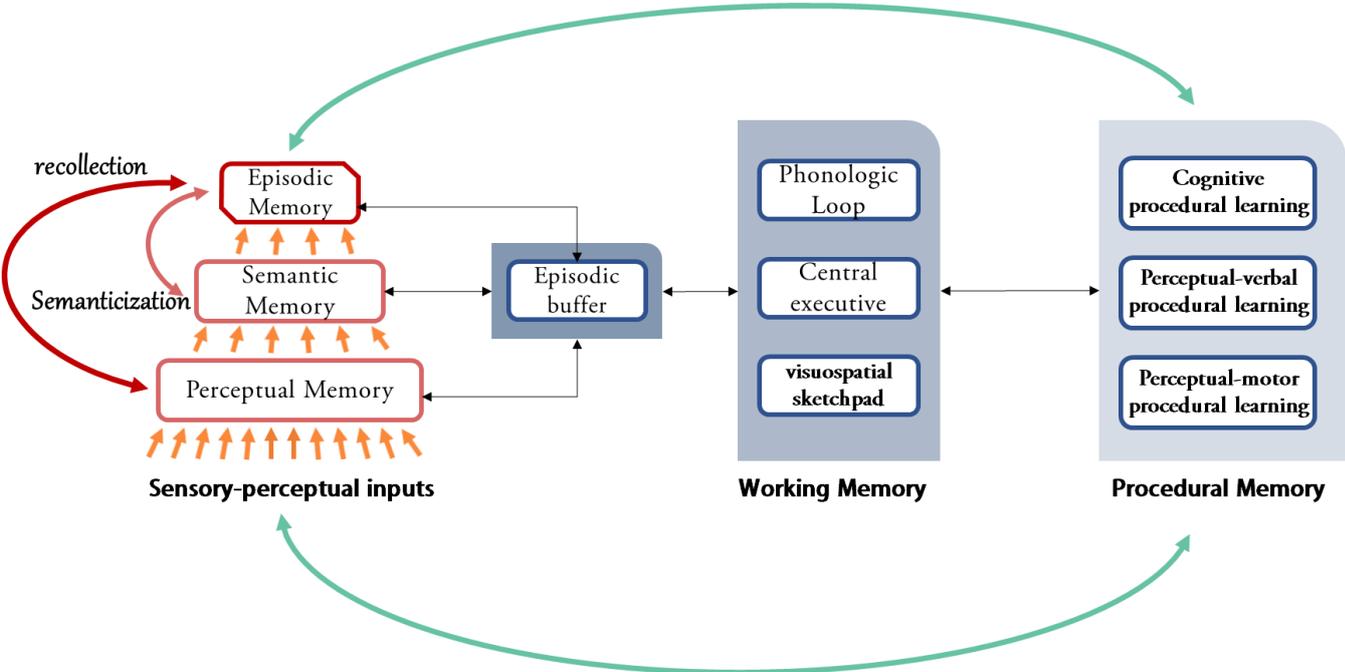


Figure 5: The MNESIS model: an integrative model of memory

This model integrates long-term declarative and procedural memory to the working memory model from of Baddeley and Hitch. It highlights the reconstructive and changing nature of episodic memory with the arrows representing the recollection and semanticization processes, but also all its interaction with other memory systems. This model is of great interest because it points out the interactive nature of memory, that do not act in isolation but with other systems. (Adapted and translated with permission via Copyright Clearance Center: Eustache and Desgranges, 2008; Eustache et al., 2016; Copyright © 2008 Springer Nature & Copyright © 2019 Elsevier).

2.1.2 *The Autobiographical framework*

Martin Conway described the memory of personal past events in the broader framework of AM, with the **Self Memory System (SMS)** (Conway and Pleydell-Pearce, 2000; Conway *et al.*, 2019; Figure 6). He presented AM as conceptually based. All autobiographical memories are constructed on a semantic framework that is defined according to current goals, affective states, mood, knowledge, and external environment. The SMS model precisely describes different AM *representations*, dividing them into two types: autobiographical knowledge and episodic memories.

Autobiographical knowledge contains semantic information about the self, goals, others, activities, locations, evaluations, *etc.* These knowledges are hierarchically structured according to their level of specificity, from the most abstract and conceptual to the most representational and event-specific (Conway and Pleydell-Pearce, 2000): *Life story schema* contains knowledge about the self that covers the life span; *Life-time periods* contain knowledge over shorter period of life, for example, the college years; *General events* concern even shorter period of time, such as repeated or long-lasting events with for example, every Christmas, or a three-day festival. According to Martin Conway, **episodic memories** are composed of a frame associated to several episodic elements. The frame is a semantic basis on which can be added these elements, which are defined as being non-verbal and “sensory-perceptual-affective” (Conway, 2009). Episodic memory is the most experienced-near representation. All these AM representations, from the most conceptual to the most experienced-near interact together when a memory is retrieved. The SMS model explains that memory can be retrieved at each level of specificity, the general knowledge memory being the most accessible (Conway, 2005). The encounter with a memory cue will result in the activation of a component in one of those AM’s structure. The activation will spread along the links existing between structures until the creation of a transitory mental map. This spreading activation is under the control of a control system that takes current goals into account (Conway, 2005; Conway and Pleydell-Pearce, 2000; Conway *et al.*, 2019).

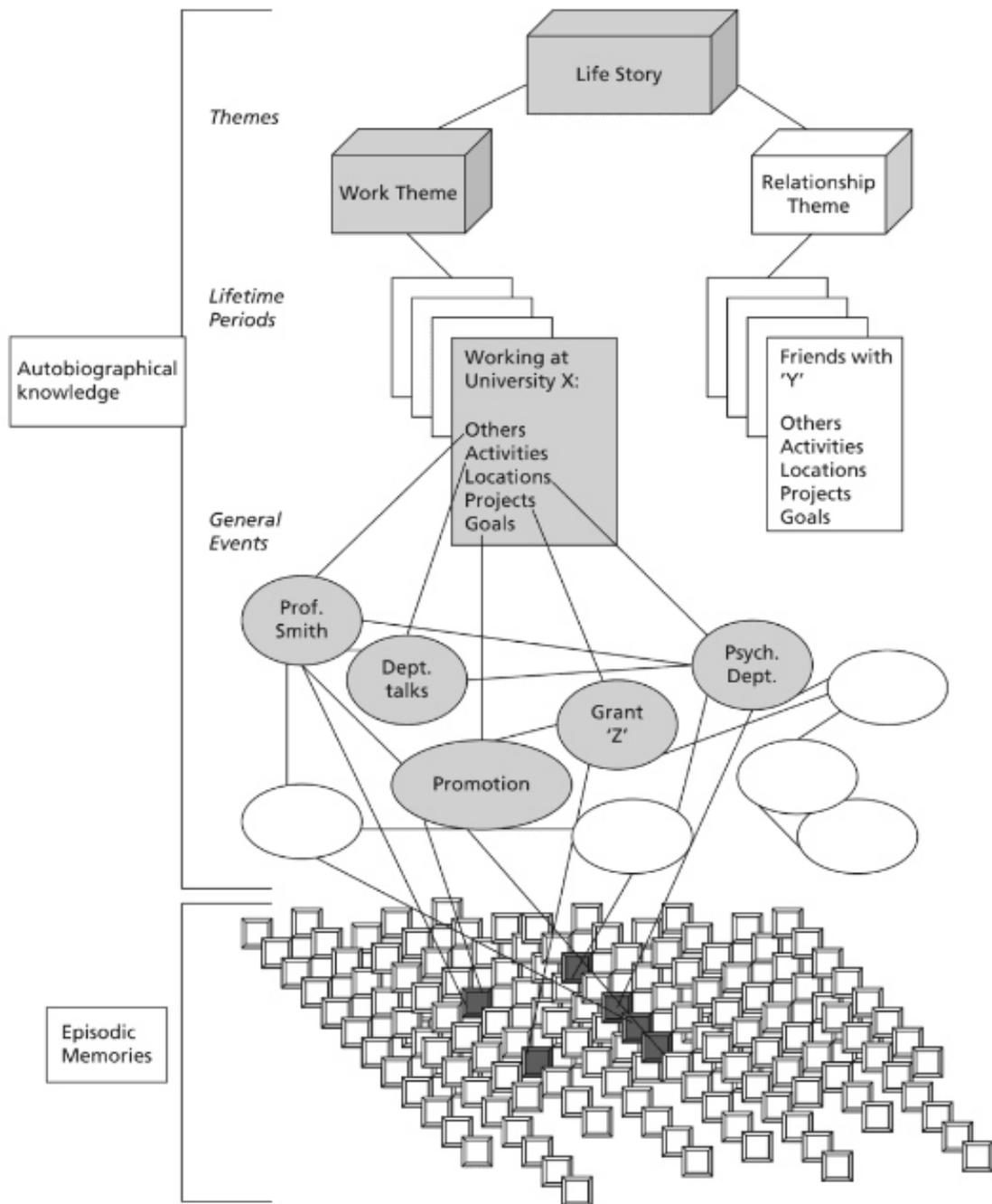


Figure 6: The self-memory system

According to Martin Conway, AM are constructed based on autobiographical knowledge and episodic memories. Autobiographical knowledge is hierarchically structured according to the level of memory specificity. Episodic memories represent to most specific and experience-near level of event memory. (Reproduced with permission: Conway, 2005; Conway et al., 2019; Copyright © 2005 by Elsevier via Copyright Clearance Center & Copyright © 2019 by Oxford Publishing Limited (Academic))

2.1.3 Present definition of episodic memory

The definitions of EM proposed by Endel Tulving and Martin Conway are not very different from each other. According to Martin Conway EM can be understood as a specific AM, involving both a semantic/conceptual frame and “sensory-perceptual-affective” information. The definition of Endel Tulving describes a “sensory-perceptual-affective” memory that can be evoked separately from its semantic frame. However, both systems lie on different hypotheses, focus on different aspects of EM, and rely on different paradigms. For example, Endel Tulving explains that EM is the most ontogenetically and phylogenetically evolved memory system, while Martin Conway explains that this is the case uniquely if SM is taken into account (for a review comparing both frameworks, see: Picard *et al.*, 2009).

These two approaches give clues to understand how the memory that allows the recollection of past events works and is about. It is a highly associative and thus a multidimensional memory. Retrieved events are about the self, they are autobiographical in nature, can be replaced in chronological order, and the retrieval is associated (in humans) with a sense of self. They are thus the keeper of an individual’s life story. Many life events are available to be retrieved, and among them the majority are weakly accessible to consciousness. However, even these weakly accessible events may have an effect on individual’s personality and uniqueness (Conway, 2009).

The retrieval process of EM has been described as similar to, or close to imagination processes (Schacter and Addis, 2007; Schacter *et al.*, 2012). Indeed, the remembering of an event implies its reconstruction: an event is not remembered as an unchanged story, parts of the event are unconsciously put together to form a story which is coherent with the current state and goals of the rememberer. Thus, the retrieval concerns both the reconstruction of past events or the imagination of future events using encoded and stored episodic information (Conway *et al.*, 2019). As remembered events are not exact copies of what has truly happened (they are reconstructed), they are not veridical and can be subjected to mistakes. The proposed mechanism of this phenomenon is that when memories are retrieved, they are put in their first labile form. Retrieved memories become thus more susceptible to behavioral and physiological interferences (Scully *et al.*, 2017). This dynamic nature of EM is highly relevant taken the possible roles of EM/AM. It is goal directed (Conway *et al.*, 2019): the past is recollected to sustain the present, and the future is imagined with the aim to predict and prepare to future situation.

For the next parts of the present manuscript, we decided not to restrict the bibliographic context to the Endel Tulving’s or the Martin Conway’s frameworks. We will thus report bibliographic works on both EM and AM, and will refer to **memories of personal events**. However, as both

frameworks rely on distinct methodologies and paradigms, we will first see what these paradigms are.

2.2 Experimental approaches

The memory of personal past events can be studied with two different approaches. *Laboratory approaches* focus on the investigation of core EM processes through measures of performance and veracity, in a controlled fashion. *Autobiographical approaches* focus on the investigation of events encoded naturally in everyday life through the analyses of their content at retrieval. However, more and more researchers tend to use paradigms that mix both laboratory and autobiographical approaches, called *laboratory-ecological approaches*, to study EM in a more natural way. In the next section, an overview of these three kinds of paradigms will be given, first the laboratory, the ecological, and finally the laboratory-ecological approaches.

2.2.1 Laboratory approach: Episodic memories studies

Historically, memory of personal events has been studied using verbal stimuli with a list-learning approach, just as Hermann Ebbinghaus did. In those experiments, participants learn a list of words during an **encoding phase**. Then, after a retention time their memory is tested during a **test phase**. Endel Tulving's work was inspired by this classical paradigm implementation, and he constructed his first definition of EM on that basis (Staniloiu *et al.*, 2020). Each step of memory (encoding, retention and test/retrieval, see: 1.1) is defined and monitored by the experimenter (Tulving, 1995), defining the *laboratory approaches*. In these paradigms, researchers measure modification of performance after changing one or more conditions at one memory step. Even though this is far from real conditions of remembering, these experiments were highly useful to reveal factors that affect different aspects of EM.

2.2.1.1 The memory of the cue: From free recall to recognition

Free recall and recognition are the main methods allowing the investigation of EM of the cue. These methods rely on the same principle as presented before, with an encoding and a test phase that are highly controlled, the main difference lying on retrieval type. In *free recall* paradigms, participants are asked to report at recall as many encoded items as possible. Experimenters then measure the mean number or the proportion of recalled items, they analyze the recall order, items that are missed or *intrusions* (recalled items that were not in the encoded list). Free recall paradigm was mainly useful to study participants' strategies during recall, or to understand errors that are made. For example, the Deese-Roediger-McDermott (**DRM**) paradigm had been conceived based on free recall to specifically study recall errors. Experimenters create lists in which all words refer to another one, which is not in the list. For example, a list can contain "monoi – deckchair – sea – sun – summer – sand", all related to

the word “beach”, which is not in the list but will later be recalled by most participants. As free recall is a difficult task, recall can be cued with the first letter of a word, a learning-associated word, or word’s category (“emotion” category for the word “happy”). *Cued recall* paradigms were shown to be less associated to EM processes than is free recall (Tulving, 1995), and was more chosen in studies analyzing simple associative memory, and pro- or retro-active interferences. **Proactive interference** occurs when previously learned information impairs the encoding of new information. For example, if participants are told to first learn a list **A** in which there is the association “Maxime-brother”, and then to learn a second list **B** in which the association changes to be “Maxime-colleague”, so the proactive interference appears when “brother” is reported when they are asked to retrieve list **B** with the cue word “Maxime”. **Retroactive interference** happened when the learning of new information impairs the recall of previously learned material. In the previous example, it can be the report of “colleague” after being cued with “Maxime”, although participants are told to report items from list **A**.

Retrieving an event is easier in a non-production task, such as the *recognition paradigm*, also widely used to study EM. In these tasks in the encoding phase, a list of items is presented to participants. These items are called **target** items or **cues**. Then in the test phase, targets are presented along with **distractors**, namely items that had not been presented in the first list. Here, after a delay, participants are asked if they have already encountered those items previously in the experiment or not. This task has many variants. For example, in *forced-choice recognition* tasks many items (generally 3) are presented to the participants at test, only one being a target. Participants have then to select the target item among the distractors. This paradigm can be useful to study recognition errors, playing with discriminability between presented items at test. The most common task is the *Yes-old/No-new recognition task*. During test, targets and distractors are presented one by one, and participants must choose to answer either “Yes, I recognize that item”, or “No, I do not recognize that item”. From the experimental conditions (target vs. distractor) and the participants' behavioral responses (“Yes” vs. “No”), four response categories were defined: **Hit** and **Miss** occurred when the target items were accurately recognized or incorrectly rejected, respectively, and correct rejection (**CR**) and false alarm (**FA**) occurred when the distractor items were correctly rejected or incorrectly recognized, respectively. The report of only correct answers hides a part of participant’s behavior. Aiming to distinguish participants who have 100% Hit because their recognition memory was good, from participants who just responded “Yes” all the time, the signal detection theory (**SDT**) framework can be applied to data analyses (Tanner and Swets, 1954). According to SDT, the decision to give a “Yes” or a “No” response is described as relying on memory strength: the more a stimulus seems familiar, the more participants will answer “Yes”, and the decision is made according to a criterion. The SDT allows thus the establishment of a

discriminability index, usually called d' , and a criterion such as the response bias measure referred to as c (Besson *et al.*, 2012). D' metric is a performance score, as it informs about participants' ability to discriminate targets from distractors. The criterion c changes according to participants' tendency to give a "Yes" or "No" answer whatever the stimulus type is. The computed value c informs thus about participants' behavior. If its value is below zero, that means that participants have given more "No" than "Yes" answers. Their behavior is then labelled as being "conservative". At the opposite, with a value higher than zero, the participants' behavior is labelled as being "liberal" because more "Yes" than "No" answers were given. Finally, a value close to zero informs that participants are bias free.

2.2.1.2 The memory of the cue: Remember/Know and Source memory

The memory of the cue can rely on different processes. It is now commonly accepted that recognition responses are given following two processes: **familiarity**, a fast and automatic process involving perceptual and semantic memory, and **recollection**, a process requiring an extensive memory search and the retrieval of the context of encoding (Besson *et al.*, 2012; Yonelinas and Jacoby, 2012; Yonelinas *et al.*, 2010).

Endel Tulving set up the Remember/Known (R/K) paradigm (Tulving, 1985) to study processes that are associated to recognition. In this paradigm, he applied his theory about consciousness states associated with different memory system to the recognition paradigm. He made the hypothesis that a stimulus can be recognized either with a familiarity or a recollection process. Participants may **know** or **remember** that the presented item is a target. A *know* response involves that participants give their answer using a familiarity judgment that relies on SM and that is associated with a noetic state of consciousness. A *remember* response involves that participants have recollected the whole encoding experience; it relies on EM and is associated with an auto-noetic state of consciousness. Thus, when participants questioned their subjective feeling and decide whether their recognition was known or remembered, this phenomenological report should help to disentangle processes behind their recognition abilities.

Source memory paradigms provide a way to disentangle in an objective way recognition based on familiarity and on recollection of the encoding context. In a typical source-memory paradigm, items are encoded with two or more different sources. The "source" can have multiple meanings (Mitchell and Johnson, 2009): it can refer to a place on a screen (right-left, or on each corner of a square or a picture), different voices, colors, or a temporal sequence. At test, items are presented to participants in a neutral way, for example in the center of the screen, in black, or with a new voice. Participants then had to say if the item has been previously encoded and to specify its encoding context. This paradigm has been shown to be

efficient to assess EM: the associative nature of the task matches with EM's associative nature, and amnesic patients showed impaired ability in source memory (Gold *et al.*, 2006; Shimamura *et al.*, 2006). However, EM requires the binding of at least three information: the object (the focus) of the event, its location, and its associated occasion, while source memory paradigm measures the binding between two information only, and thus can be retrieved either by a recollection process or by simply knowing the association. Moreover, neural substrates that underpins source-memory and R/K paradigms had been shown to be only partially overlapping between each-others (Frithsen and Miller, 2014). Thus, the source-memory task does not completely allow for the study of EM in its globality.

2.2.1.3 Drawbacks

All tasks presented above were designed as ways to unravel mechanisms underlying EM. Free recall and recognition focus on performance about item memory, whereas R/K and source memory paradigms focus on performance on associations between an item and associated phenomenological traits or behavior. They all tap into different core aspects of EM and are related to different neural substrates (Cheke and Clayton, 2013; Frithsen and Miller, 2014; Stevenson *et al.*, 2020), thus allowing together to have a more general idea of EM mechanisms. But that said, they all allow to study only a small part of the whole process underlying memory of our past events. Only Tulving's R/K paradigm investigates one phenomenological aspect, and paradigms investigating their associative nature involves only a simple association. Moreover, reconstructing an EM involves higher level processes allowing for the selection of the most relevant information in a pool of stored events' memory according to the remembering context and current goals of the rememberer.

2.2.2 Ecological approach: Autobiographic memories studies

Memory of events is also investigated in real life contexts using approaches that focus on AM content and associated phenomenological traits. The hypothesis is that accuracy is less functionally relevant than the content of the memories. Memories are dynamic and changing, susceptible to interferences because the environment is also changing, providing perceivers the ability to adapt to an always changing environment. As AM is complex in levels of specificity and global structure (Conway, 2009; Conway *et al.*, 2019), numerous ways of investigating AM had been conceived, assessing life episodes, global periods or life story. Here, we will focus on AM tests that study memories of life episodes, that are triggered either voluntarily or not.

2.2.2.1 Voluntary AM

2.2.2.1.1 Cueing methods

Created by Galton in 1879 and later developed by Crovitz and Schiffman in 1974⁴, the cueing method is now the most widely used method in AM studies (for example, see Barzykowski *et al.*, 2019; Bonnici and Maguire, 2018). In this paradigm, a list of cues, generally words, are provided to participants, who must tell the first personal memory that pops up and that is related to the cue. Variants of this method exist: participants could be asked to specify a date to their memory, and cues can change from words to other types of stimuli, such as sounds or odors. It is also possible to ask participants to remember specific types of events: the most vivid (memories that are intense, rich, clear, detailed), important ones, happier, sadder, earlier memories, *etc.* In the *event cueing variation*, a series of first remembered memories serves as cues to a second series. This allows for example to describe relations between recalled memories and to make inferences about the construction and organization of AM store.

2.2.2.1.2 Interview methods

The *Autobiographical Interview* is a widely used semi-structured interview developed by Levine and colleagues (Levine *et al.*, 2002). This method requires participants to tell one memory for five lifetime periods: early childhood, teenage years, early adulthood, middle age, and the previous year. This method is used to encourage participants to give detailed memories. For each recall, two probes are given: a general one and then a specific one. General probes consist of a general question designed to encourage a more detailed recall (“Can you tell me more about...”). Specific probes are a kind of structure interview that is conducted to elicit even more details. It consists of standardized questions that are grouped into five categories: event, time, time integration, place, other sensory information and emotion. The Recall data are then transcribed into a text that is segmented into units. These units are then categorized depending on their type: *internal* details, referring to SM or *external* details, referring to EM. Then, information in each unit is divided in narrower categories (such as time, emotion, thoughts...) that allows details to be rated in terms of episodic richness.

The TEMPau method, the French acronym of Episodic Test of Past autobiographical Memory, is a semi-structured interview that have been set up at the same time by Pascale Piolino and colleagues (Piolino *et al.*, 2009). This test transfers theories of EM coming from laboratory experiments into an interview constructed on the same basis as the Autobiographical Interview. Participants have to recall four memories from each of five periods of their life. As these memories must be episodic, a detailed instruction about the nature of event to recall is given to the participants. Recalled events must be unique and include specific places and

⁴ Referenced in the Handbook of Research Methods in Human Memory (2018)

dates, and the event duration must be less than one day. Then, after each recall, participants are required to judge their memory as having been recollected, knew, or made up with a guess (uncertain memories), and whether they recalled their memory with an observer (3rd person) or a field (1st person) perspective. Along the task, participants were helped with standardized general prompts. Information was checked on a subsequent test or by asking family members.

2.2.2.1.3 *Towards more verifiable methods*

All methods presented in the previous sections (2.2.1 and 2.2.2) have the drawback to be poorly verifiable. In TEMPau efforts are made in this direction but the check with family members also relies on their memory and can thus be also distorted or the event can be forgotten. The diary recording method have been set up to overstep this drawback, while staying close to ecological conditions. On a given period, participants are asked to record events of their daily life. Thanks to recent technological advances, some studies used a portable camera, which takes several pictures in a day to provide memory cues. Later, diary entries and pictures are used to cue participants memories with the Galton cueing method (see for a review: Dubourg *et al.*, 2016).

Another method involves the creation of realistic and controlled events in the laboratory, through exhibition or specific professional procedures, as for example a mandatory training for hospital staff (Diamond *et al.*, 2020a, 2020b). These events can then be evaluated with common AM tests.

2.2.2.2 *Involuntary autobiographical memories*

As humans can have at least as many voluntary than involuntary triggered memories (Rubin and Berntsen, 2009), methods have been set up to study involuntary AM. The easiest way is to ask participant to keep a diary in which they report each time a memory pops up in their mind involuntarily. Clear instructions are generally given about the nature of memory experimenters are interested in, and participants are usually requested to answer a few specific questions when the memory recall occurs, and then later on the same day (*i.e.*, see: Berntsen, 1996). These questions depend on the aim of the study and can be about the type of cue that triggered the recall, or phenomenological properties.

Involuntary AM can also be measured in a more controlled way, as when the recall takes place in the laboratory (Barzykowski and Niedźwieńska, 2018). For example, participants can be engaged in a task where they must find a specific pattern in a picture full of different ones. Then, sometimes, a cue is presented without participants being aware of it, and they must report each time an AM is recalled (Barzykowski and Staugaard, 2016).

2.2.2.3 Drawbacks

Numerous paradigms have been conceived to unravel different aspects of AM. As they are constructed to be representative of a global and ecological AM, they do not allow to reach the exactness and precision that can be conducted under laboratory studies. Moreover, those methods do not allow for the study of AM errors, and the information about what have been forgotten, what have been misreported, distorted, or falsely incorporated into the memory report is of great interest to understand how memory works. Finally, as these methods rely all on verbal report, they do not give access to memories that are difficult to put into words.

2.2.3 Laboratory-ecological episodic memory approaches

Laboratory and ecological approaches are complementary and bring together important information that improve our understanding of EM processes. However, they both have drawbacks. Ecological approaches do not allow for the study of specific encoding mechanisms. Laboratory approaches are too artificial, which deprives them of rich phenomenological characteristics, and are too simplistic, as they did not involve highly associative memories. Moreover, these two EM testing approaches rely on both common and distinct neural bases that may each inform us about part of the underlying EM processes (Cabeza *et al.*, 2004; Chen *et al.*, 2017; McDermott *et al.*, 2009; Monge *et al.*, 2018; Roediger and McDermott, 2013). Nowadays, more and more researchers are going toward more complex laboratory approaches, closer to real life conditions. These tasks are called laboratory-ecological tasks.

2.2.3.1 The development of laboratory-ecological methods

Encouraged by declarations stating that non-human animals cannot have an EM, Nicola Clayton and colleagues set up a task based on the 1972's Endel Tulving definition of EM, that matches his example: "*Most, if not all, episodic memory claims a person makes can be translated into the form: "I did such and such, in such and such a place, at such and such a time."*" (Clayton and Dickinson, 1998; Tulving, 1972). In their task, they take advantage of the natural behavior of scrub jays to hide their food, and their preference to eat wax worms over peanuts, two kinds of food that become uneatable either quickly (wax worms) or not (peanut). Hiding them for long or short delays in specific places, they demonstrated the ability of scrub jays to have an integrated and complex memory. Scrub jays could indeed recall where eatable food (a wax worm or a peanut) can be found depending on the delay and their preference. They thus created the first *What-Where-When (WWW)* paradigm in which the *What* represents the focus of the event, the *Where* its spatial location, and the *When* its relative time. Madeline Eacott and Alexander Easton suggest to use a broader definition of the temporal condition, adapting the *What-Where-When* to a *What-Where-Which (context or occasion) paradigm* (Eacott and Easton, 2010; Easton and Eacott, 2008). These tasks have been widely used in

animal research, allowing for the demonstration of an episodic-like memory in rodents, non-human primates, dolphins, dogs, horses, insects and cuttlefish (for review, see: Madan, 2020).

Laboratory-ecological approach of EM has also been used in human studies. In these tasks, participants typically explore different complex environments at encoding, which are distinct from each other according to their spatio-temporal context. Remembering occurs differently depending on the protocol. For example, participants can be asked to freely recall encoding events, their answers accuracy being verified with other questions (Holland and Smulders, 2011; Plancher *et al.*, 2013). In some other studies, the retrieval phase is more controlled, requiring participants to select the correct spatial location and the right occasion among previously explored ones (Easton *et al.*, 2012; Saive *et al.*, 2013, 2014a, 2015). For example, Anne-Lise Saive and colleagues used a *WWW* paradigm. They asked participants to smell odors (*What*) localized at different places (*Where*) in a specific contextual image (*Which*) for three days. Each day, novel odors were presented in distinct places of a new background image. Thus, several specific What-Where-Which / Odor-Place-Context combinations were encoded. On the fourth day, previously smelled odors were presented among an equal number of distractor odors. Participants had to select the odor they have smelled previously, and then specify in what place and in what image they were associated. Despite its difficulty, the participants succeeded in the task above chance, and their correct complete retrieval was associated to more remember than know responses (Saive *et al.*, 2015), a result that is in agreement with another similar study using abstract visual forms (Easton *et al.*, 2012).

2.2.3.2 The contribution of virtual reality

All methods in section 2.2 could benefit from the development and the popularization of a promising new tool in EM research: virtual reality (VR). Virtual reality involves the 3D simulation of one or more environments in which participants are immersed, and with which they can interact in real-time. This definition involves many types of VR devices: it can be computer-controlled (desktop-VR), dispensed through head-set apparatus, or through the use of a simulator (Smith, 2019). The immersion, namely the degree to which the participants are “captured” by the virtual environment and can forget the real world around them, can vary across devices. Immersivity also varies according to features that have been constructed and included in the environment as for example the number of visual details, the availability of multimodal information, and the possibility to freely move and interact in a deep way with the environment (Smith, 2019).

Using VR in EM allows laboratory tasks to be based on richer encoding conditions, that are closer to reality than the study of word lists, while keeping a control on all memory stages (Foudil *et al.*, 2021). Indeed, VR allows for the creation of flexible and complex environments

that can be lived in the field perspective by the participants, and in which they can move freely to explore them in a similar way as in real life. As this tool offers many advantages, more and more studies are using VR to study EM, where VR have been shown to be efficient to study EM in aging, neurodegenerative disease, and neurocognitive rehabilitation (La Corte *et al.*, 2019; Plancher and Piolino, 2017).

2.3 Neural bases of memory of personal event

To ensure the complex and reconstructive ability allowing the recollection of personal events, numerous brain areas are involved. Among these areas, the MTL plays a central role, allowing cortical activity pattern drawn during the encoding of an event to be stored as episodic traces (Renoult *et al.*, 2019). With the MTL, cortical brain regions organized in different networks are commonly activated. Altogether, they allow a cue that can be provided internally or externally (in the environment) to reinstate episodic traces, to recreate a complex mental representation, and to relive it with phenomenological experiences and a sense of self.

2.3.1 Medial temporal lobe

Numerous clinical studies from the 18th century until now, including William Scoville and Brenda Milner's pioneer research describing the case of H.M. patient and other patients, have shown the central role of Hipp and surrounding regions – the MTL – to encode and restore recent declarative memories (Dickerson and Eichenbaum, 2010; Scoville and Milner, 1957; Squire and Zola-Morgan, 1991). The MTL includes several brain regions that all have been shown to be involved in the same global function, the support of declarative memory systems. It is composed of the Hipp, the entorhinal (**EC**), the perirhinal (**PRC**) and para-hippocampal (**PHC**) cortices. The Hipp receives inputs from the EC, that receives itself information from the neocortex through the PRC and the PHC (Squire and Zola-Morgan, 1991). Although seemingly forming an entity, all regions of the MTL have specificities. According to the Binding of Items in Context theory (**BIC**) proposed by Howard Eichenbaum and colleagues (Figure 7), information coming to MTL is of two types: the object of the event (the *What* component), and the context (the *Which context* component) that makes this event unique (Dickerson and Eichenbaum, 2010; Reagh and Ranganath, 2018). These two types of information come from parallel pathways (Ritchey *et al.*, 2015) to be integrated in the Hipp. The Hipp then binds this information to form an episodic memory trace (Dickerson and Eichenbaum, 2010; Poppenk *et al.*, 2013; Reagh and Ranganath, 2018; Ritchey and Cooper, 2020; Ritchey *et al.*, 2015). Information forming the *What* component comes mostly from unimodal associative areas and reach PRC through the ventral visual pathway. Processed information from PRC is then conveyed to the lateral entorhinal cortex (LEC) that sends most of it in the anterior part of the Hipp. PRC and anterior Hipp are involved in fine-grained conceptual processing, such as the association between non-spatial stimuli and between a stimulus and its features, including

semantic ones, thus allowing for the identification of an item. Coarse grained spatial and contextual information that are experienced-based and perceptual comes mostly from plurimodal areas through the dorsal visual pathway and are processed by the PHC. This contextual information is then predominantly sent to the posterior part of the Hipp through the medial entorhinal cortex (**MEC**).

The Hipp is a highly integrative area that is anatomically a convergence point of all associative areas through integration of afferences coming from PHC/MEC and PRC/LEC. Indeed, stimulus information coming from PRC/ LEC and contextual information coming from PHC/ MEC are both sent in dentate gyrus, then the CA3, and CA1, a pathway known as the trisynaptic circuit (Dickerson and Eichenbaum, 2010). The information goes then out of the Hipp through the subiculum which send a feedback to neocortical areas. The Hipp is thus able to create a representation of this association between items and scene.

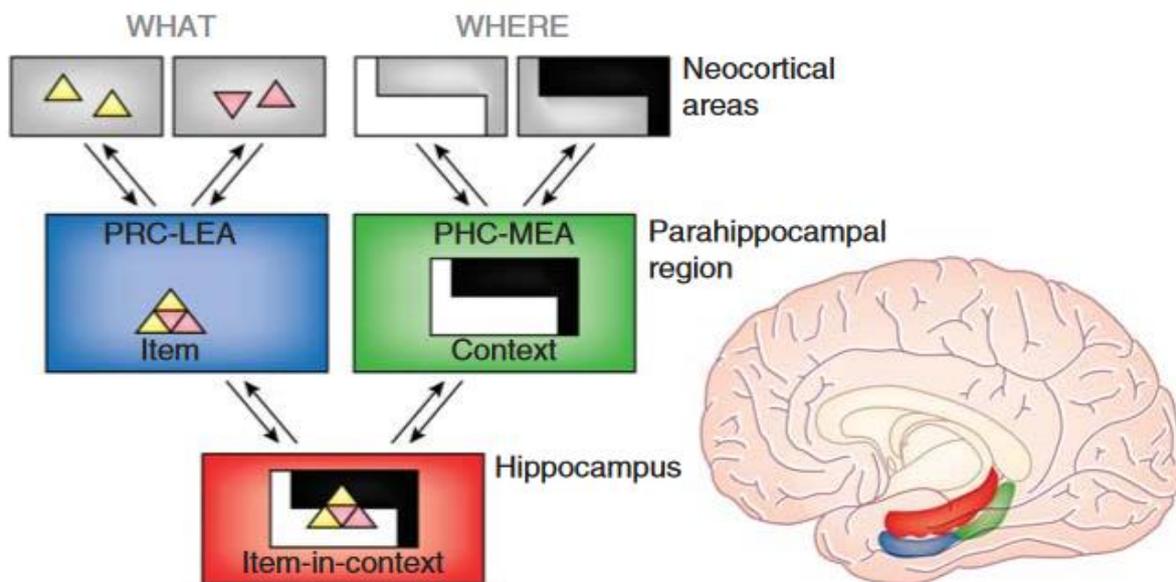


Figure 7: Schema of the Binding of Items in Context (BIC) model

This model represents the episodic information processing in the medial temporal lobe. Episodic information from the neocortex is processed in two pathways according to their type. In the medial temporal lobe, information concerning the item (the What) is processed by the parahippocampal cortex and medial entorhinal area (PHC and MEA) before reaching the hippocampus. The information concerning the item location (the Where) is processed by the perirhinal cortex and lateral entorhinal area (PRC and LEA) before reaching the hippocampus. The hippocampus is then able to integrate the information from the What and the Where pathways, and to send back processed information to the neocortical areas. (Reproduced with permission: Dickerson and Eichenbaum, 2010; Eichenbaum et al., 2007; Copyright © 2007 by Elsevier & Copyright © 2009 by Springer Nature)

2.3.2 Cortical pathways

The ability to build a mental representation of personal events also requires the involvement of cortical areas. Recollecting an event involves its reconstruction (Renoult *et al.*, 2019; Ritchey and Cooper, 2020), which requires the access and the representation of separated event features and their relation with each other, a self-perspective, self and conceptual knowledges, a capacity to relive the restored event through mental images, and the involvement of monitoring and evaluation processes that will guide the retrieval decision (Dickerson and Eichenbaum, 2010; Ritchey and Cooper, 2020). Moreover, events are lived, encoded, and reconstructed based on existing schema, namely the dynamic and abstract representations that have been acquired following repeated experiences (Renoult *et al.*, 2019). For example, a schema of a birthday cake is that it has candle on it.

Numerous studies have shown the importance of a *core recollection network*, engaged during AM tasks and also during event imagination (Renoult *et al.*, 2019). This network involves cortical areas that tend to be left lateralized, probably due to the involvement of semantic processes: medial prefrontal cortex (**mPFC**), lateral prefrontal cortex (**lPFC**), angular gyrus (**AG**), posterior cingulate cortex (**PCC**), retrosplenial cortex (**RSC**), and other areas that allow for visuo-spatial imagery with for example cuneus and precuneus (Cabeza and St Jacques, 2007; Dickerson and Eichenbaum, 2010; Rugg and Vilberg, 2013; Svoboda *et al.*, 2006). According to Peggy St Jacques and colleagues (2011), these cortical areas can be separated based on their role into four networks working together to allow for recollection, the connection between those networks being modulated by recollection strength. The first two networks would be implicated in both construction and elaboration phases during retrieval: 1) the mPFC network would be involved in the initiation and maintenance of recollection. Notably, during recollection, dorsal mPFC is involved in self-referential processing and ventral mPFC is thought to monitor memory processes and to be responsible of the feeling of rightness associated to the retrieved memory representation, and 2) the MTL network, as presented earlier, allows the reconstruction of mental scenes and contextual information. The two remaining networks have been presented to be involved in the construction phase, during retrieval: 3) the left lateralized frontoparietal is responsible of cognitive control process and strategic search, and the 4) cingulo-opercular networks is involved in the maintenance of goals throughout the memory construction.

Other studies have reported a segregation of cortical areas involved in recollection based on information quality. For example, the gradient theory proposed that information in cortico-hippocampal systems is organized with an antero-posterior axis (Poppenk *et al.*, 2013; Ritchey *et al.*, 2015; Sheldon *et al.*, 2019), as presented in Figure 8. Precise, fine grained information would rely on posterior regions while coarse grained, gist-like, conceptual (semantic)

information (that mainly concerns the central element of the episode) would rely on anterior regions (Dickerson and Eichenbaum, 2010; Ritchey and Cooper, 2020; Sheldon *et al.*, 2019). Maureen Ritchey, Charan Ranganath and colleagues (Reagh and Ranganath, 2018; Ritchey and Cooper, 2020; Ritchey *et al.*, 2015) proposed a model in which two global networks are working together to access, construct and elaborate personal event memories: the anterior temporal system and the posterior medial system. Anterior temporal network gathers regions that are functionally connected to the anterior Hipp: Amygdala, lateral orbitofrontal cortex (OFC), PRC, anterior fusiform gyrus, anterior inferior temporal cortex. Its proposed role is the encoding, long term storage and retrieval of items and their perceptual, semantic, emotional associated information in the form conceptual gist-like (coarse-grain) information. The posterior medial network is based on several regions that process contextual, fine-grained information. These regions are functionally connected to posterior Hipp: PHC, RSC, PCC, precuneus, AG, anterior thalamus, pre-subiculum, mammillary bodies, mPFC, parietal cortex. Its main role in is to encode events' context as spatial, temporal and causal relations between items, to store them as situation information, and to reconstruct it in a self-perspective at retrieval. Thus, the anterior temporal system and the posterior medial system work together to allow for the reconstruction and the recollection of memories, the posterior medial network providing the contextual scaffold in which core event information coming from the anterior temporal network can be included.

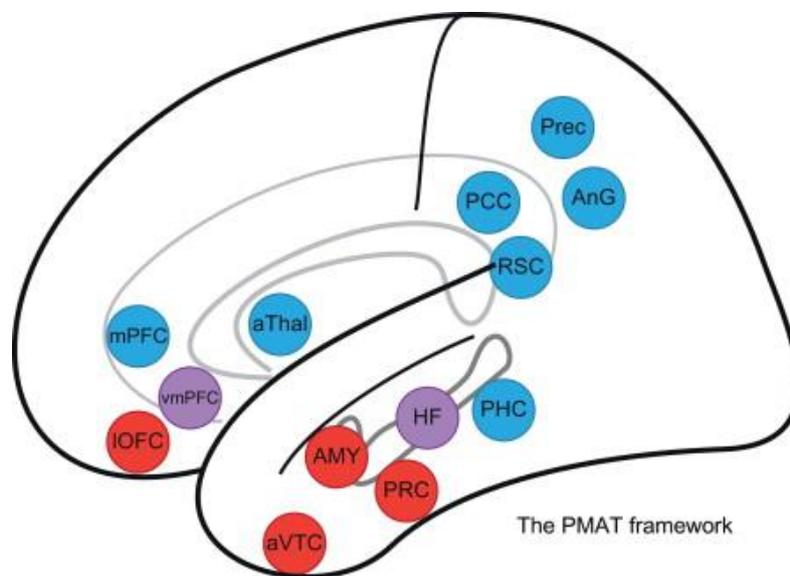


Figure 8: Posterior Medial – Anterior Temporal (PMAT) framework

The PMAT framework represents the neocortical processing of personal event memory, with the PM network (in blue) processing contextual information whereas the AT network (in red) processes the item information. These two networks converge in two areas, the hippocampal formation (HF) and the ventromedial prefrontal cortex (vmPFC). Regions of the PM system include the PHC, RSC, posterior cingulate (PCC), angular gyrus (AnG), precuneus (Prec), anterior thalamus and mammillary bodies (aThal), and medial prefrontal cortex (mPFC). Regions of the AT system include the perirhinal cortex (PRC), anterior ventral temporal cortex (aVTC), amygdala (AMY), and lateral orbitofrontal cortex (IOFC). (Reproduced with permission via Copyright Clearance Center: Ritchey *et al.*, 2015 ; Copyright © 2015 by Elsevier.)

2.3.3 Influence of the experimental approach on neural bases of personal events

As seen above (2.2), the memory of past events can be studied with different investigation approaches. Laboratory studies restrict the complex memories of past events to simpler, basic elements that can be easily controlled. They involve a poor encoding context and are based on high monitoring capacities. Episodes are often recent and meaningless. In contrast, autobiographical tasks study real-life memories but are lacking experimental control. They involve a rich encoding context, they are based on extrapolation that requires a high elaboration phase to complete missing information of episodes that are older, self-centered, and meaningful and remembered with a high reliance on mental imagery (Chen *et al.*, 2017; McDermott *et al.*, 2009; Renoult and Rugg, 2020). These two kinds of paradigms may thus rely on partially different cerebral bases. Numerous studies agreed with that statement. First, if laboratory and ecological tasks deal with the same mechanisms, we could think that people showing high capability on one task also show the same abilities on the other one. However, people showing highly superior AM that can accurately remember specific events of their past, show only average performance on recognition paradigms (Roediger and McDermott, 2013).

Regarding cerebral bases, the comparison of recognition tasks (laboratory measure of EM) to the cued paradigm (or Galton-Crovitz paradigm, ecological measure of AM) have shown they rely on different neural networks and on brain regions that are only weakly overlapping (Chen *et al.*, 2017; McDermott *et al.*, 2009; Roediger and McDermott, 2013). Laboratory tasks are shown to rely on regions that are responsible of high monitoring (anterior fronto-parietal network) and the perception of novelty/familiarity (Parietal memory network, Chen *et al.*, 2017). Zachary Monge and collaborators (2018) extend these findings by comparing laboratory (free recall) and ecological (AM, cued paradigm) tasks in both search and recovery phases during the retrieval. Although no differences were shown in the search phase, a cognitive control area (superior parietal lobule) was more activated in the laboratory task, whereas regions of the Default Mode Network (**DMN**) were more activated in the ecological task during the recovery phase. These regions are known to be involved in internally directed cognition that occurs during rest and though to be involved in the creation of mental models (Ritchey and Cooper, 2020). So, it was shown that ecological recall of AM involves a more bilateral activation (while more pronounced to the left hemisphere) of regions belonging to the DMN (Chen *et al.*, 2017; McDermott *et al.*, 2009).

3 Memory and the senses - How sensory modalities influence memory of personal events

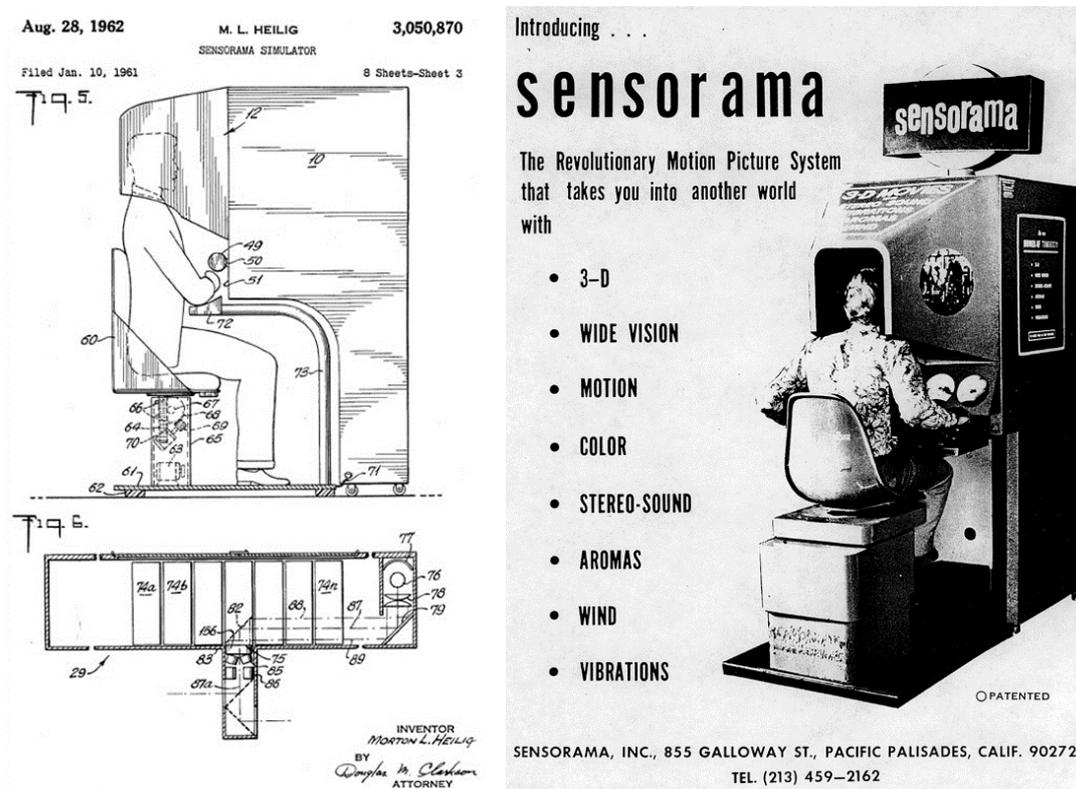


Figure 9: The Sensorama Simulator (Heilig, 1962)

The sensorama simulator is an old invention that can be presented as a multisensory television set. The “sensospectators” are meant to live a registered “movie” as it occurs in real life, using most of their senses and not just two, as in movie theaters or in most of memory research.

Sensory modalities shape our daily life. The loss of one of them, even the one we think the most “useless” can have bad consequences on quality of life, mental and/or physical health. Each sensory modality allows humans to perceive, memorize and recall the world in diverse ways, allowing humans to have complex and rich experience, as Morton Heilig tempted to reproduce using the *sensorama simulator* (Figure 9). Each sensory modality brings complementary information to understand and adapt to the current or future environment. This adaptation behavior also explains the flexible nature of human memories. Thus, memorizing and recalling personal events through these different “lenses”, brings humans best route to adapt to its environment. Unfortunately, the contribution to each sense to the memory of our personal events is poorly known. Thus, in the next chapter the effect of three types of stimuli that are important in our daily life are presented: Face (3.1), Music (3.2) and Odor (3.3). Each of these parts are constructed the same way. The first section presents the importance of such stimulation in daily human life (3.X.1), follows a section concerning their behavioral and neural perceptual characteristics (3.X.2), and finally a section that presents behavioral and neural bases of the memory of the stimuli and of the memories evoked by the stimuli (3.X.3).

3.1 The case of face

3.1.1 Importance of faces in human everyday life⁵

Faces have a considerable importance in human everyday lives. Human faces convey a lot of information that are relevant for the perceiver. They provide information about an individual's origin, gender, familiarity, identity, emotion, intention, and can be judged on attractiveness and dominance. They are thus relevant, for example, to detect individuals belonging to our peers, to know how to react and how to interact when meeting someone, and for the reproductive behavior.

Human's ability to process faces is heritable (Wilmer *et al.*, 2010) and varies across people, going from very limited abilities to very developed ones (Burton *et al.*, 2010; Russell *et al.*, 2009; Stantic *et al.*, 2021). The reduced ability to identify faces can have negative consequences on human psychological health. Interindividual analysis, for example, showed that low scores at face identification tests is correlated with social anxiety traits. Face perception and recognition is also often impaired in various pathologies, *i.e.*, autism disorder, Turner's syndrome, schizophrenia, and Alzheimer's and Parkinson's diseases. The selective deficiency in face recognition is called *Prosopagnosia*. It can be congenial or acquired. Particularly, congenital prosopagnosia is very informative regarding its influence in everyday life, as it does not extend to other impairments (Corrow *et al.*, 2016). Not having the ability to understand who are in front of them, humans suffering from prosopagnosia are afraid to offend their relatives, and colleagues, leading to traumatic social interaction, chronic anxiety, loss of self-confidence and a restricted social circle. To identify people, patients suffering of prosopagnosia scheduled who they will probably meet in a day, and must rely on distinctives features, such as hairstyle, gait, voice or usual clothing style. So, any change in all of these features can lead them to not identify themselves or a family member (Corrow *et al.*, 2016; Yardley *et al.*, 2008), although this ability to recognize familiar faces with different haircut, hair color or tan, in a variety of different contexts seem incredibly easy to other people. Here are two examples taken from the study of Lucy Yardley and her collaborators (2008): "*My sister said to my mother that I just looked straight through her, as if she were a stranger—but of course to me it was as if she was a stranger.*" (participant 24), or "*It's hard to make a strong connection with anybody, it takes a lot of meetings ... years before you can really get to know somebody, so you become a bit excluded*" (participant 08).

⁵ The review of Olivier Pascalis and David J. Kelly (2009) guided the construction of the following part.

3.1.2 Face perception, emotion and cognition

3.1.2.1 Characteristics

The importance of faces in social and everyday life and their omnipresence leads some researchers to sustain that humans are experts in face processing (Young and Burton, 2018a). However, this notion of expertise is debated and had been presented to be reliable only for familiar faces (Rossion, 2018; Sunday and Gauthier, 2018; Young and Burton, 2018a, 2018b, 2018c). The processing of familiar and unfamiliar faces seems to rely on different mechanisms. Unfamiliar faces are thought to be processed on a more pictorial basis than familiar faces. This explains for example that matching the identity of a face presented in a picture to another picture of that face but taken on another occasion, embedded in a set of novel faces' pictures, is more difficult for an unfamiliar face than for a familiar face (Figure 10 A and B). Another evidence is that pictures of familiar faces are better recognized than unfamiliar faces if emotion or view is changed (Bruce, 1982). As we choose to study the influence of unfamiliar stimuli to control the previous experience of participants with stimuli and to render them comparable across sensory modalities, this chapter will be devoted to unfamiliar face processing. Thus, from now on, the word "face" will be referring to unfamiliar faces, unless explicitly stated.

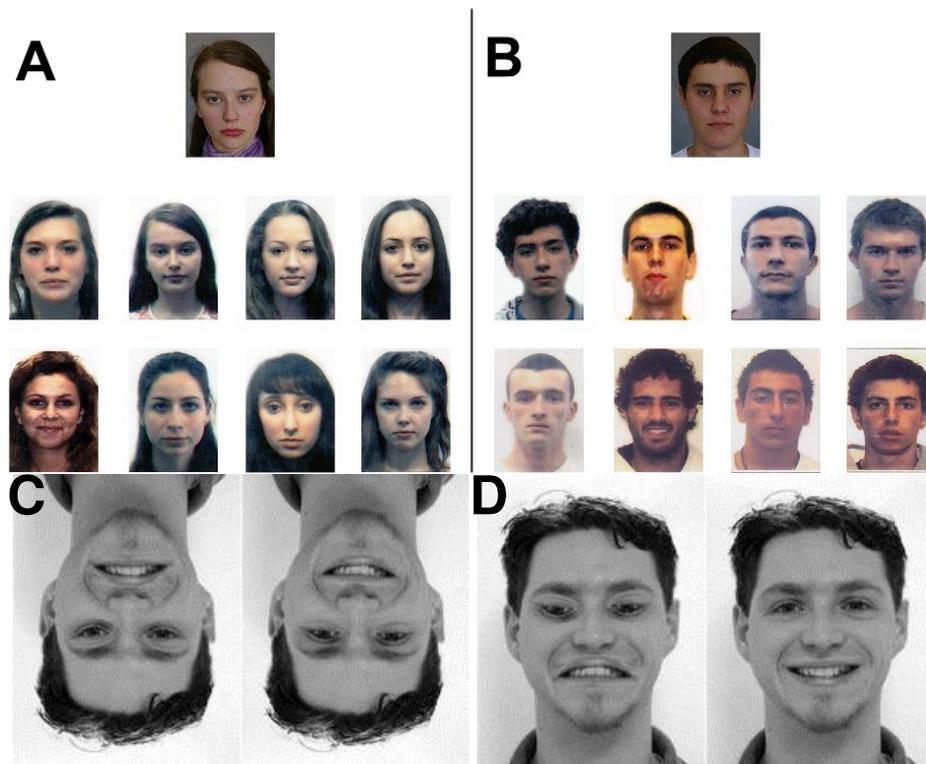


Figure 10: Unfamiliar face perception

A. and B. Example of two one-to-eight face matching arrays. In this task participants have to decide whether the target in the top image is present in the gallery below and then to select it⁶; C. and D. Thatcherized faces of R. Le Grand seem to be more normal when inverted than upright, demonstrating the thatcher effect. (Reproduced with permission. A. and B.: White et al., 2015, CC BY 4. ; C. and D.: Maurer et al., 2002 Copyright © 2002 by Elsevier).

⁶ *Answer: A. The face of the girl is on the top row, second from the left, B. The face of the boy is not present.*

Humans are particularly responsive to faces. This sensitivity arises early in development, as newborns show a preference to look at faces rather than at other stimuli only a few hours after birth, they prefer to see their mother's face than other faces after few days, and infants are thought to be able to categorize faces (Pascalis and Kelly, 2009). In adults, detecting a face is fast: humans take approximately 140 ms to do a saccade toward a face, with the fastest saccade being at 100 ms (Crouzet *et al.*, 2010). Compared to other visual stimuli, such as airplanes, animals, butterflies and car, faces are detected more rapidly, even for patients with prosopagnosia.

The easiness and speed with which faces are detected and identified may be due to their processing. Daphne Maurer and collaborators (2002) formalized the first stages of face processing. They explained that faces undergo three types of *configural processing*, the process that allows the perception of the relations among faces' features. The first step is the *detection of the first-order relations*, that is, the detection of the common basis on which all faces are constructed: two eyes above one nose and one mouth, forming a T-shaped configuration. Then, faces undergo a *holistic processing*. Face features are integrated in a way that explains that a face is perceived as a whole, instead of the sum of all its different parts. The third step concerns the sensitivity to *second order relations*: the distance between face's features that contributes to the singularization of different faces. Faces also undergo a *featural processing*, also called *analytic processing*, referring to the fact that faces can be also discriminated efficiently thanks to the shape, color or luminance of its features (Yovel and Belin, 2013). The evidence of a configural processing that distinguishes faces from other visual objects, is attributed to the face-inversion effect. This effect shows that faces are better and more quickly perceived and recognized when presented upright than inverted, whereas it does not impair other visual objects processing, nor featural information processing for faces (Maurer *et al.*, 2002; Taubert *et al.*, 2011; Yin, 1969). The example of inverted Thatcherized faces (*Figure 10*), that are constructed by inverting the eyes and the mouth on an inverted face (the first one being the face of Margaret Thatcher), is enlightening. With this manipulation, faces seem to be more normal when inverted than upright (Thompson, 1980). This effect was explained by the fact that when inverted, featural processing operated more quickly than configural processing, and not for upright faces. This may cause a change of processing importance from configural to featural processing. When inverted, the fact that eyes and mouth are inverted is more important than the global look of the face. At the opposite, when this face is presented upright, the global look is more important, and thus a feeling of bizarreness is felt (Carbon and Leder, 2005).

During early processing steps, faces are categorized according to their social group (gender, origin, age, occupation, sexual orientation, *etc.*) or emotional expressions. Face categories are constructed on a gist-based memory of the features that are shared by all members of a group and that are stored as categories' prototypes. In practice, when a face is encountered, its features are quickly matched with these prototypes, allowing their categorization (Rapcsak, 2013; Rapcsak and Edmonds, 2011). With this rapid categorization, social category is shown to affect early face processing. For example, ingroup faces are processed more holistically and are more subjected to individuation (emphasizing features that distinguish a face from other ones) than outgroup faces (Hugenberg *et al.*, 2011). Face emotions are thought to be expressed similarly across cultures and peoples, thanks to a combination of stereotyped configuration of muscle movements that are resumed by the Facial Action Coding System (FACS; Ekman and Friesen, 1976; Matsumoto *et al.*, 2008), while this point of view is not universal (Barrett *et al.*, 2019). The recognition of emotional expressions relies on different processes. A rapid process allows the perception of emotional categories based on low-spatial frequency visual features (the broad, general configuration of face elements). A second, longer process, allows the perception of high-frequency spatial features, which capture thinner facial details (Blair, 2003).

A specific class of faces, that has become more and more important with the virtualization of life, is virtual faces. These faces are thought to be processed similarly to other faces. They are sensitive to inversion (Balas and Pacella, 2015; Kätsyri, 2018), and emotional expression seems to be conveyed similarly between real and virtual faces (Dyck *et al.*, 2008). However, even when matched for fine-grained visual features (contrast, brightness, color), virtual faces are accurately discriminated from real faces (Kätsyri, 2018). Also, inversion effects seem to act to a lesser extent for virtual than real faces, leading to the conclusion that they are processed as an out-group face category (Balas and Pacella, 2015; Kätsyri, 2018).

3.1.2.2 Neural bases of face processing⁷

Figure 11 illustrates the neural bases of visual objects processing. As perception of all visual stimuli, face perception begins with the integration of light information that is reflected from a face by the photoreceptors located in the retina (Figure 11A). The information is transmitted through the bipolar cells, and then the ganglion cells, which axons form the optical tract. The signal coming from the optical tract is then separated into several pathways. Around 90% of the axons go to the lateral geniculate nucleus of the thalamus before reaching the primary visual cortex, located in the occipital cortex. The 10% remaining axons go through other

⁷ Chapters on visual perception and object recognition of the “*Cognitive Neuroscience, The Biology of the Mind*” book (Gazzaniga *et al.*, 2019) guided the construction of this section.

structures. Among them the superior colliculus is thought to be involved in a pathway that recruits the pulvinar nucleus.

Information coming from the primary visual cortex is then processed in other visual cortices also located in the occipital lobe, allowing for the representation of the stimulus to be more and more complex. The information is represented in the geniculate nucleus and the visual cortex following a retinotopic map: information that is spatially close from each other in the seen environment is represented by neurons that are close from each other's in the cortex.

From the primary visual cortex, visual information is separated into two pathways Figure 11B. The ventral stream (occipito-temporal or parvocellular pathway) is thought to represent the visual identity and ends in the inferior temporal cortex. The dorsal stream (occipitoparietal or magnocellular pathway) is thought to be devoted to the visual location and ends in the posterior parietal cortex. Crucially, the frontal cortex influences the bottom-up information coming from the occipito-temporal pathway. It generates predictions about the visual scene, that influence information coming from lower areas to allow for faster processing of that visual information.

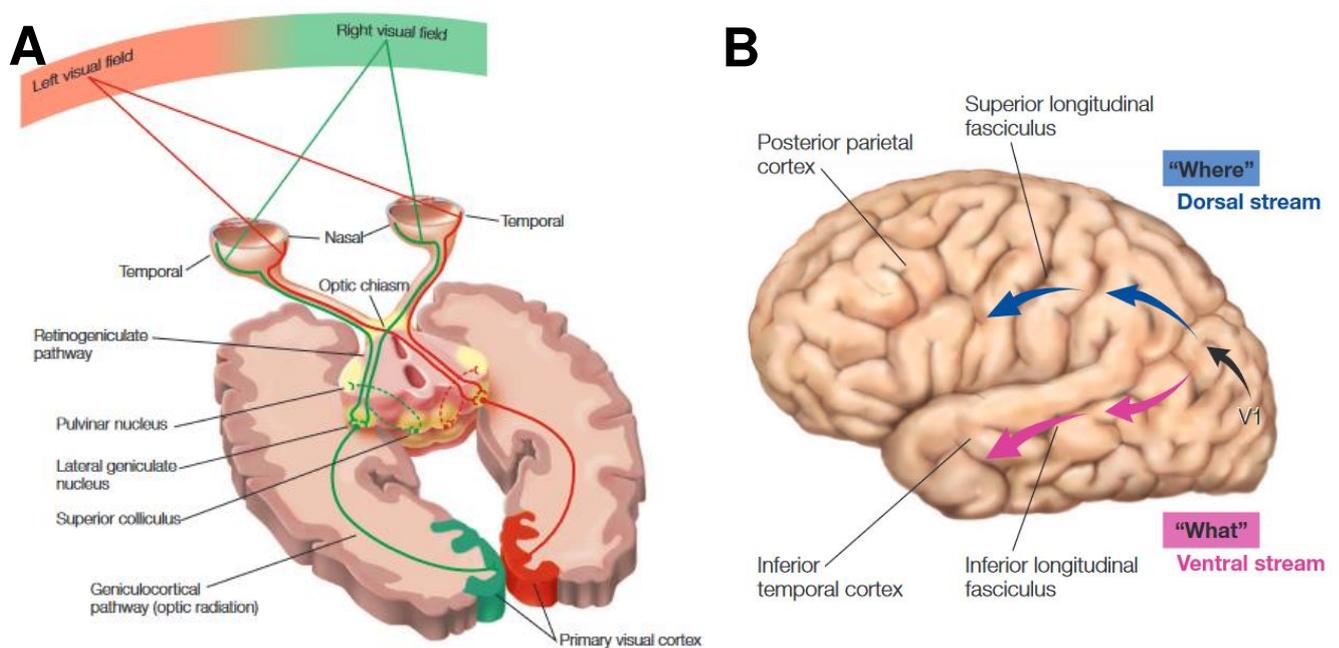


Figure 11: The visual pathway, from the eyes to neocortical areas

A. Visual projection to the primary visual cortex. B. Visual Where and What pathways. (Rights: Gazzaniga et al., 2019; chapter 5.6).

Figure 12 represents the neural bases of faces. Faces are processed in what is called face-selective areas, in other words, cortical areas that are more activated in response to faces than to other visual objects. These areas, that have been extensively studied since the seminal work by Nancy Kanwisher and collaborators (1997), are mainly located in the right hemisphere along the ventral visual stream (for reviews, see: Duchaine and Yovel, 2015; Haxby *et al.*, 2000; Rapcsak, 2019). Face processing, whether faces are real or virtual (Kegel *et al.*, 2020), involves the contribution of two networks: a core network, made of ventral selective-face areas, and an extended network, made of dorsal-selective face areas. The core ventral network allows for the modality-specific visual analysis of faces and the processing that goes from the detection of the *first-order relations* to their identification. It is located in the occipital face area, the fusiform face area, and the anterior temporal lobe face area (Figure 12b). These regions process information hierarchically, from a more structural and view-specific processing located in posterior regions, to a more abstract and view-independent processing located in the more anterior regions. Thus, the occipital face area, presented as the entry point of the network, conducts a more structural encoding, processing the facial features that are view-dependent, whereas anterior temporal face area processes view invariant features in a more holistic fashion. The extended dorsal network involves the posterior superior temporal sulcus (STS) face area, the anterior STS face area, and the inferior frontal gyrus (Figure 12a). These regions are thought to be devoted to the dynamic processing of faces and are constantly interacting with the ventral face processing pathway. While the core network only processes visually centered information, the extended network process multimodal information. It allows the recall of associated semantic, emotional, and episodic information of familiar and/or unfamiliar faces and thus relies on related areas.

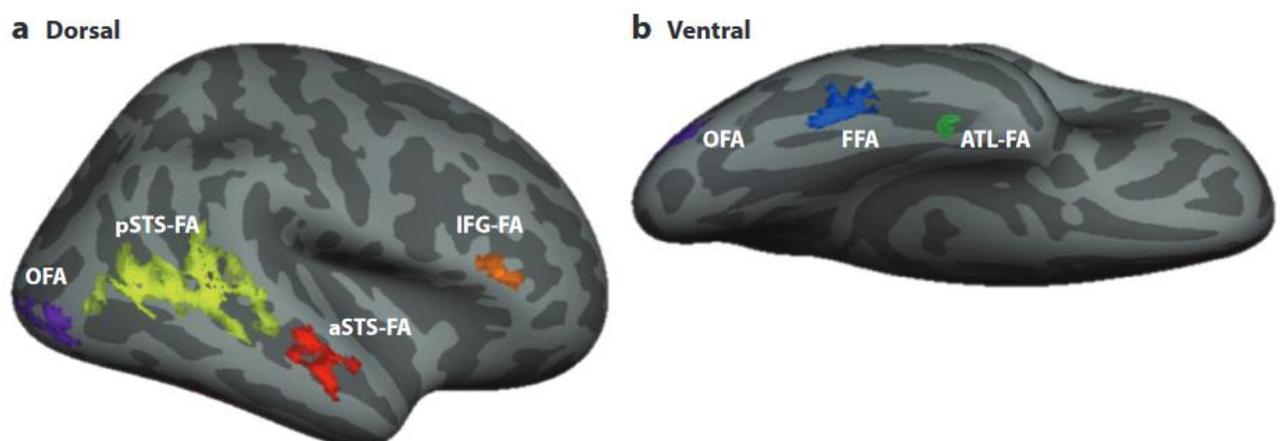


Figure 12: Face-selective areas

a. Dorsal face area: the posterior superior temporal sulcus face area (pSTS-FA), the anterior superior temporal sulcus face area (aSTS-FA), and inferior frontal gyrus face area (IFG-FA); and b. Ventral face processing areas. the occipital face area (OFA), the fusiform face area (FFA), and the anterior temporal lobe face area (ATL-FA). (Reproduced with permission: Duchaine and Yovel, 2015; Copyright © 2015 by Annual Reviews, Inc.)

Facial emotional expressions are processed by two pathways (for review, see: Johnson, 2005). A fast subcortical route allows low-spatial-frequency information to be conveyed from the colliculus and the pulvinar to the amygdala via the magnocellular channel. Its activation modulates the activity of face-selective cortical areas (see Vuilleumier and Pourtois, 2007 review for further evidence). This route explains the fast detection of coarse stimulus features that indicates, for example, a potential threat. The second pathway, a slower route, is conveyed via parvocellular channels to the face-selective cortical areas and is tuned to process high-spatial-frequency information. It allows processing fine-grained information, such as expression-related wrinkles. Emotional expressions then activate face-selective cortical areas, in particular the fusiform face area (Vuilleumier and Pourtois, 2007), and regions that are also activated for visual scenes, such as the amygdala, OFC, mPFC, inferior temporal cortex, and extrastriate occipital cortex (see the meta-analyses by Dean Sabatinelli and collaborators, 2011). When real faces are compared with virtual moving faces, fear emotion elicits similar activation in ventral temporal region and amygdala, whereas dorsal temporal region and inferior frontal gyrus are more activated for real faces (Kegel *et al.*, 2020). This result may explain that coarse emotion information accurately conveyed by virtual faces, whereas their dynamic features may be insufficiently similar to real faces to activate dorsal face processing pathway.

3.1.3 Face and memory

3.1.3.1 Memory for faces

Recognizing specific pictures of faces seems to be incredibly easy, relatively to other objects. Notably, the recognition score is higher for pictures of faces than for pictures of scenes (Sato and Yoshikawa, 2013), houses, airplanes, moving persons (Yin, 1969), odors (Cornell Karnekull *et al.*, 2015), and voices (Cortes *et al.*, 2017; Damjanovic and Hanley, 2007). This recognition of unfamiliar faces is picture-based, in contrast to familiar faces which recognition is identity based (Longmore *et al.*, 2008). It is indeed easier to reject (decide that it is a new face) an unseen picture of a face seen at the encoding (same identity, different picture) for an unfamiliar face than for a familiar face. However, it is easier to recognize the identity of a face seen at the encoding but transformed at recall (same identity, different picture) for a familiar face than for an unfamiliar face (Armann *et al.*, 2016).

Several factors impair unfamiliar face recognition (see for review: Johnston and Edmonds, 2009). The own-ethnicity bias (also called the other-race effect) explains better face recognition for people with who we used to live comparing to people from other origin (Feingold, 1914). A similar effect is observed for other social categories, such as gender

(Wright and Sladden, 2003), and age (Mukudi and Hills, 2019). As virtual faces are processed as out-group faces, such as other age, gender or ethnicity, recognition is also impaired in comparison to real faces (Balas and Pacella, 2015), leading to more FA scores in recognition paradigm (Kätsyri, 2018). Face recognition is also impaired when faces wear a face-mask (Freud *et al.*, 2020), when the view or the expressed emotion is changed between study and test (Bruce, 1982; Liu *et al.*, 2014) and, comparing to other visual objects, face recognition is more affected when the stimulus is inverted between study and test (Yin, 1969).

The effect of emotional face expression on face recognition has been mostly studied by two kinds of paradigms. (1) In the first paradigm, the expressed emotion is the same at encoding and retrieval (Cortes *et al.*, 2017; Johansson *et al.*, 2004; Keightley *et al.*, 2011; Wang, 2013). Negative faces seem to be better recognized than neutral ones (Keightley *et al.*, 2011; Wang, 2013), even if sometimes no emotion effect is shown (Johansson *et al.*, 2004) or neutral stimuli allow for better recognition (Cortes *et al.*, 2017). This inconsistency may be linked to the reduced number of studies that investigate participant's subjective feeling relative to the displayed emotional stimuli. In the studies cited above, only one has measured face pleasantness and evaluate its influence on memory scores (Wang, 2013) and shows that negative faces, which had been rated as being less pleasant than positive and neutral ones, are better recognized than neutral ones. To note, positive face recognition was not different from neutral ones. (2) In the second paradigm, while emotional faces are presented at encoding, they are presented in their neutral form at retrieval (D'Argembeau and Van der Linden, 2007; Liu *et al.*, 2014; Righi *et al.*, 2012; Shimamura *et al.*, 2006), or the reverse (Chen *et al.*, 2015). This paradigm allows studying the importance of emotion on the recognition of face identity at diverse memory stages. Increased identity recognition rate are observed for happy faces relative to other displayed emotions (Chen *et al.*, 2015; D'Argembeau and Van der Linden, 2007; Shimamura *et al.*, 2006; but see: Liu *et al.*, 2014; Righi *et al.*, 2012). This paradigm lies on identity recognition, as participant must recognize the identity of the face despite a change of emotion, which have been shown to impair recognition (Bruce, 1982). Moreover, such a manipulation is infeasible for other sensory modalities. Thus, the first paradigm will be of interest in this PhD research.

3.1.3.2 Memories evoked by faces

Irrespective of face familiarity, few studies have investigated the specificity of faces as a memory cue. Faces are associated with more recollection responses in a R/K paradigm than are voices (Brédart and Barsics, 2012a; Damjanovic and Hanley, 2007). Faces are also able to evoke more AM, and these AM contain more details than those evoked by music stimuli (Belfi *et al.*, 2016). Face-evoked AM are associated with a stronger feeling of being brought

back in time than AM evoked by a song or its lyrics (Cady *et al.*, 2008). Regarding the effect of emotion, the same inconsistency holds as in recognition paradigm, with negative emotion giving more (Johansson *et al.*, 2004; Wang, 2013) or less (Cortes *et al.*, 2017) remember responses than neutral stimuli.

3.1.3.3 Neural basis of face memory

Using fMRI, Iris Trinkler and collaborators (2009) compared the encoding and recognition of unfamiliar, famous, and personally known faces. Two cortical networks are revealed. The first network underpins the recognition of previously seen faces irrespectively of their familiarity (Hit versus CR). It is composed of bilateral middle cingular cortex, PCC, precuneus, and ventral striatum, without any hippocampal activation. The absence of hippocampal activation is consistent with the fact that a selective damage in Hipp area does not affect face recognition (Bird and Burgess, 2008). The second network is associated to the contrast between personally known and famous faces with unfamiliar faces, irrespectively of the memory phase (encoding or retrieval) and score. Thus, it may sustain the familiarity associated to known faces. It is composed of bilateral RSC, the anterior parieto-occipital sulcus, angular gyrus, Hipp, amygdala and lateral temporal cortices. Noteworthy, while famous and personally known faces activate the same networks, the activation varies with the degree of familiarity. Particularly, the MTL activation was positively correlated with previous knowledge of the faces.

Regarding emotion effects on memory triggered by faces, Michelle L. Keightley and collaborators (2011) showed that, relative to neutral faces, emotion carried by faces enhances neural network activity associated to correct recognition. In particular, correctly recognized negative items, that were better recognized than positive or neutral ones, were associated with increased neural activity in amygdala, Hipp and bilateral frontal and parietal cortices.

3.2 The case of music

3.2.1 Importance of music in human everyday life

For the most of us music forms an important part of our everyday life and we would be agreeing to the well-known citation of Friedrich Nietzsche: “*Without Music, life would be a mistake*” (Nietzsche, 1908). Numerous children were exposed to music before birth and were embedded in it throughout their life, and most people can cite a musical piece that can resume some life episodes. More than the importance of music exposure in western adult population, music is universal. Music was indeed detected and understood by all human cultures investigated (Fitch and Popescu, 2019; Särkämö *et al.*, 2013; Stevens, 2012). Moreover, considering the broader notion of musicality, that is, a more natural form of music, without any social or cultural construct, music even appears in non-human species as humpback whales or some birds (Honing *et al.*, 2015). Music is thus not just “auditory cheesecake” (Pinker, 1998), a human construct that would have the sole function to be used as a hobby, to pass time or produce pleasure. Music could have a great power in many domains (Honing *et al.*, 2015; Särkämö *et al.*, 2013).

The power of music is such that many properties had been attributed to it. Whereas some properties seem to hold only on popular beliefs without scientific basis (Schulkind, 2009), others do not. For example music can be helpful in speech and language disorder training approaches, such as dyslexia (Fiveash *et al.*, 2021). Music is also used as therapy to help aphasic patients and children with autism spectrum disorder to communicate (Särkämö *et al.*, 2013; Wan and Schlaug, 2010). Music is also highly related to motion: it is produced by movement, induces dancing and moving with rhythm at all ages in all culture even when just listening (Molnar-Szakacs and Overy, 2006; Stevens, 2012). So, gait disorders that are common in Parkinson disease or appear after a stroke can be reduced by music therapies or auditory cueing (Särkämö *et al.*, 2013).

The well-known and close link between music and emotion is shared across cultures (Fritz *et al.*, 2009). It is involved in pleasure responses to music but can also be used to manage emotion in everyday live. Music is thus used to calm babies, regulate negative emotions (Taruffi and Koelsch, 2014), or in therapies to help persons with affective disorders as depression or anxiety, and to maintain well-being in older people (Särkämö *et al.*, 2013). Music is also a mean of communication and expression that is different than language, in the sense that language requires to be in interaction with others to be understood and music do not (Cross, 2014). In that way, it is a social media that is involved in group construction, cooperation and cohesion, and that helps to construct socio-cultural identity (Cross, 2012; Honing *et al.*, 2015; Särkämö *et al.*, 2013). For example, live music concerts help to construct

social connection while being isolated at home during the recent pandemic situation (Swarbrick *et al.*, 2021).

Although the condition appears not to be well-known, people can lose the sense of musicality, a disorder called amusia that prevalence is evaluated to be at about 1.5-4% of the general population. Amusia is characterized by an impairment of perception, memory and production of music (for review, see the chapter of Tillmann *et al.*, 2015). For example, amusic individuals are unable to detect wrong notes, or to recognize a familiar melody unless lyrics are provided. Most amusic individuals are in fact unable to perceive and memorize small variations of pitch, but the impairment can affect other music properties, such as timbre or rhythm, even though to a lesser extent (Särkämö *et al.*, 2013; Tillmann *et al.*, 2015). This impairment seems to affect emotion detection and recognition in music, but also in prosody (variations in speech melody, rhythm, stress, intonation; Lima *et al.*, 2016; Pralus *et al.*, 2019).

3.2.2 Music perception, emotion, and cognition

3.2.2.1 Characteristics

Music represents varied forms of structured sound sequences that are represented by many genres and sub-genres in western cultures, from the widely studied classical music, to the less investigated metal/death-core music. Musical sequences are organized into two main axes: tonal and temporal dimensions (Janata, 2015; Koelsch, 2018). The tonal dimension of music gathered the organization of pitches into intervals, melodies, chords according to scales, modes (minor/major), and keys, giving the basis of harmony and consonance/dissonance perception. The events, *i.e.*, notes or chords, that compose a musical piece can be structured in time thus defining a rhythmic pattern in the melody, which might give rise to a regular underlying pulsation, the beat. Based on this, the structure of meter with its potential hierarchy can be built. This is for example one of the differences between a waltz (with a triple meter, 1-2-3-1-2-3) and a lullaby like “*Brother John*” (in French “*Frère Jacques*”, that is constructed with a quadruple meter, 1-2-3-4-1-2-3-4). The tempo, the number of pulses per minute, gives the speed at which a song is played. Another music dimension, but that is generally not form-bearing dimension in Western tonal music, is timber. Timber gives the identity to the auditory object, *i.e.*, whether it is a violin, or a piano. Music is thus a highly complex and well-structured sequence of pitches over time. How music is structured and composed varies according to music styles and culture, making the songs humans used to hear a part of their memories.

Unlike a picture of a face, that can be explored almost instantly, music unfolds over time. Music perception requires to keep track of previously heard notes as a melody is unfolding, an ability that involves the short term memory system (Peretz and Zatorre, 2005; Särkämö *et al.*, 2013). But music perception also requires long-term memory. In the sequence of played notes C – D

– E – F (respectively *do – ré – mi – fa*, in French notation), almost everyone would expect a G (*sol*) to follow. Exposure to music of one's culture and genre leads to the extraction of the most common and basic musical structures, patterns and regularities, and their storage into long-term memory, thanks to the cognitive capacity of implicit learning. This statistical learning allows musicians and non-musicians, to expect musical events to come, with expectations being linked to musical expressivity and emotion (Huron, 2006; Pearce, 2018; Stevens, 2012; Zatorre and Salimpoor, 2013).

Emotional response to music is complex and can be conveyed by different ways. First, it is important to distinguish the emotion that the composer want to express, which can be recognized when listening to the music, and the emotion felt by the listeners (for review, see Schubert, 2013). Music has been shown to express a large amount of different emotions, some of which are commonly studied and are well recognized by listeners (sadness, happiness, anger) - even coming from different cultures (Fritz *et al.*, 2009; Stevens, 2012) - and other do not share high agreement between listeners (jealousy, pity, cruelty, eroticism) (Juslin, 2013a; Warrenburg, 2020). Expressing emotion in music relies on some common basis. For example, using minor versus major modes and/or slow *versus* fast tempo, expresses sadness *versus* joy emotion respectively (Vieillard *et al.*, 2008), and dissonant music is less positively valenced than is consonant music (Fritz *et al.*, 2009). How are these expressed emotions felt by the listeners? Patrik Juslin (2013b) proposed that music conveys emotion through several mechanisms. *Musical expectancies* hypothesis (Huron, 2006; Huron and Margulis, 2010; Meyer, 1956) explains that emotion arises when expectations are violated or confirmed, according to the amount of uncertainty in the musical piece (Cheung *et al.*, 2019). This continuous play between expectations and their resolution are presented as the major way through which music makes feel emotions to the listener. The *emotional contagion* hypothesis proposes that listeners feel the emotion conveyed by the music, but to a lesser extent (Lundqvist *et al.*, 2009; Schubert, 2013). The *Brain stem effect* arises when an event in the music attracts attention (with dissonance, loudness, longer pause...) causing a feeling of surprise or arousal. The *rhythmic entrainment* mechanism explains that music rhythm generates a modification of physiological rhythm or entrains neural oscillations. The *evaluative conditioning* comes from the associative learning between the music and an emotional associated event or object. The *visual imagery* mechanism is the mechanism by which music evokes emotional or relaxing images. The *episodic memory* mechanism explains that the felt emotion corresponds to the memory's emotion evoked by a musical piece. The *aesthetic judgment* mechanism explains that emotion is felt while evaluating music as an art, on its form and content. Feelings induced by music can be accompanied by physiological changes as skin temperature and conductance, heart rate (Fuentes-Sánchez *et al.*, 2021; Lundqvist *et al.*,

2009), that can give rise to a chill, also called a thrill (Panksepp, 1995). The variety of emotions that is conveyed by music also depends on the context in which it is heard and on the listener. For example, some people are afraid of death metal music, notably when there is scream or growl, while others truly appreciate it, enjoying the melody and feeling the energy and the power evoked by some songs, such as *The Eagle Flies Alone* from *Arch Enemy* band. Aside from the variability evoked by different musical styles, Ernest Mas-Herrero and colleagues have shown that some people are not sensitive to music besides being able to recognize emotion conveyed by music (Mas-Herrero *et al.*, 2014).

3.2.2.2 Neural bases music processing⁸

Music is a complex and structured sound, and its perception shares many processing steps with other sound perception. Figure 13 represents the neural bases of auditory perception. After entering in the outer and middle ear, the acoustic signal enters the inner ear, into the cochlea. There, the signal is dispersed as a function of its frequency, from 20 Hz (low frequencies) at its apex to 20 kHz (high frequencies) at its base, what is called a tonotopic organization and defines the pitch dimension of the sound. This auditory information is transduced into an electric signal, that is sent through the auditory nerve to the brainstem. In the brainstem, it goes through the cochlear nucleus, superior olivary complex, inferior colliculus, where the basic features of the sound (duration, intensity, frequency, and location) are processed. After a relay in the thalamus (medial geniculate nucleus), the signal is sent to the auditory cortex, the last areas of the brain that keeps a tonotopic organization.

The auditory cortex is organized hierarchically from less associative to more associative areas, with a core region surrounded by belt and parabelt regions (Zatorre and Salimpoor, 2013). The auditory cortex is responsible for the organization, bonding, construction, and maintenance of auditory objects (for review the role of on auditory cortex regions, see Janata, 2015). Further, pitch and temporal processing take two pathways, from the auditory cortex to frontal regions (Figure 13B). Pitch processing involves the interaction between the belt / parabelt region and the inferior frontal cortex (through an antero-ventral pathway), and temporal and beat processing involves the interaction between auditory cortex and dorsolateral frontal cortices postero-dorsal pathway), together with other motor areas (supplementary and pre-motor areas, cerebellum, putamen),. From these two pathways, the brain would be able to extract statistical regularities from the auditory signal and to use them to construct expectancies (for reviews, see: Janata, 2015; Zatorre and Salimpoor, 2013).

⁸ This was written with the help of two reviews, one written by Petr Janata (2015) and the second by Robert J. Zatorre and Valorie N. Salimpoor (2013).

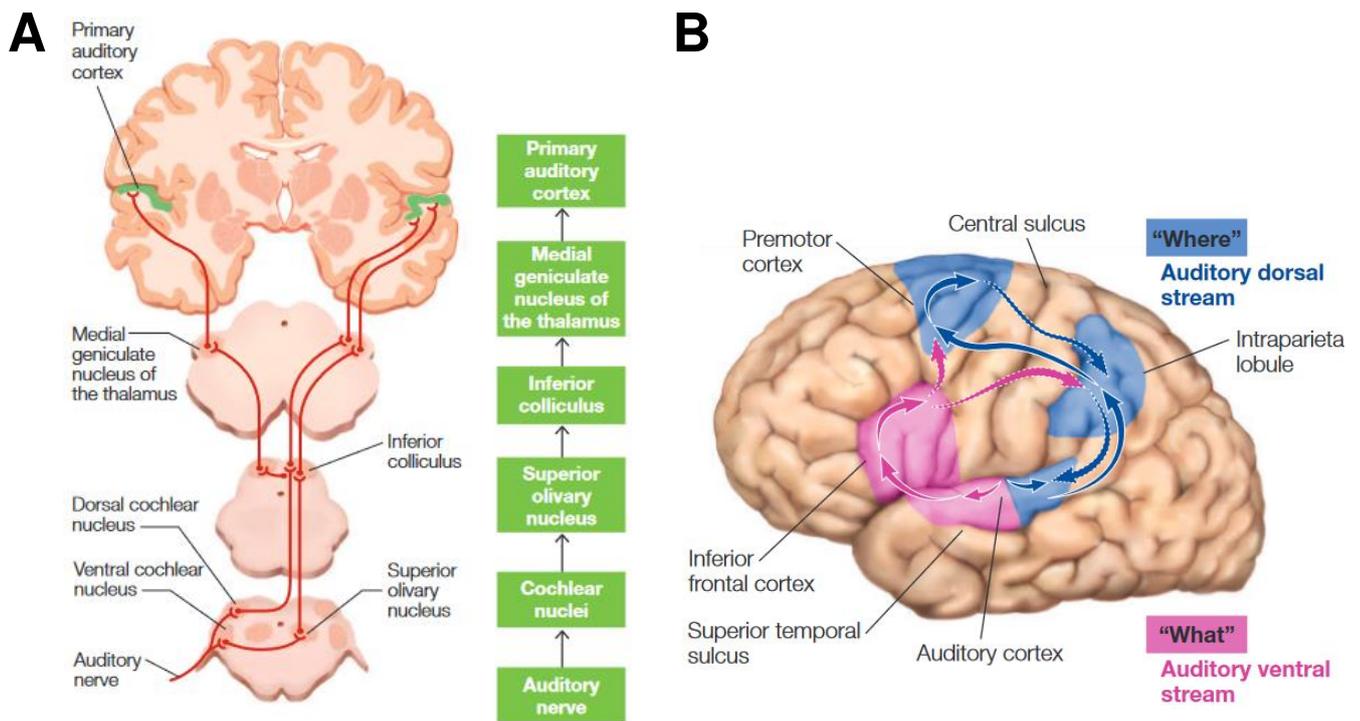


Figure 13: The auditory pathway, from the auditory nerve to neocortical areas

A. Auditory projection to the primary auditory cortex. **B.** Auditory *Where* and *What* pathways.

Rights: Gazzaniga et al., 2019 ; chapter 5.5)

The perception of simple musical features, such as music frequency or intensity involves areas described just before and involved in general auditory perception: the inferior colliculus, the auditory cortex, and the thalamus. Music processing then involves a vast amount of brain regions. In a review, Teppo Särkämö and collaborators (2013) identified five networks that could be activated by music listening, explaining complex musical features processing and the links shared between music and memory, motion and emotion (Figure 15). Higher order musical features (1) such as harmony, rhythm, and music syntax involve areas from the antero-ventral and dorsal pathways (PFC, premotor cortex, superior temporal gyrus, and inferior parietal lobe). The process that allows to keep the track of music through time (2) involves the with cingulate gyrus, prefrontal dorsolateral cortex, inferior parietal lobe. The activation of associative memories (3) involved the Hipp, medial temporal and medial parietal areas. Music is closely linked to movement (4), and via the auditory-motor coupling in the brain, music listening also creates activations in movement-related areas (Motor and somatosensory cortices, cerebellum, basal ganglia). Finally, emotion (5) is conveyed by music through the activation of striatal, limbic and frontal cortices (striatum, Amygdala, Hipp, cingulate and orbito-frontal cortices).

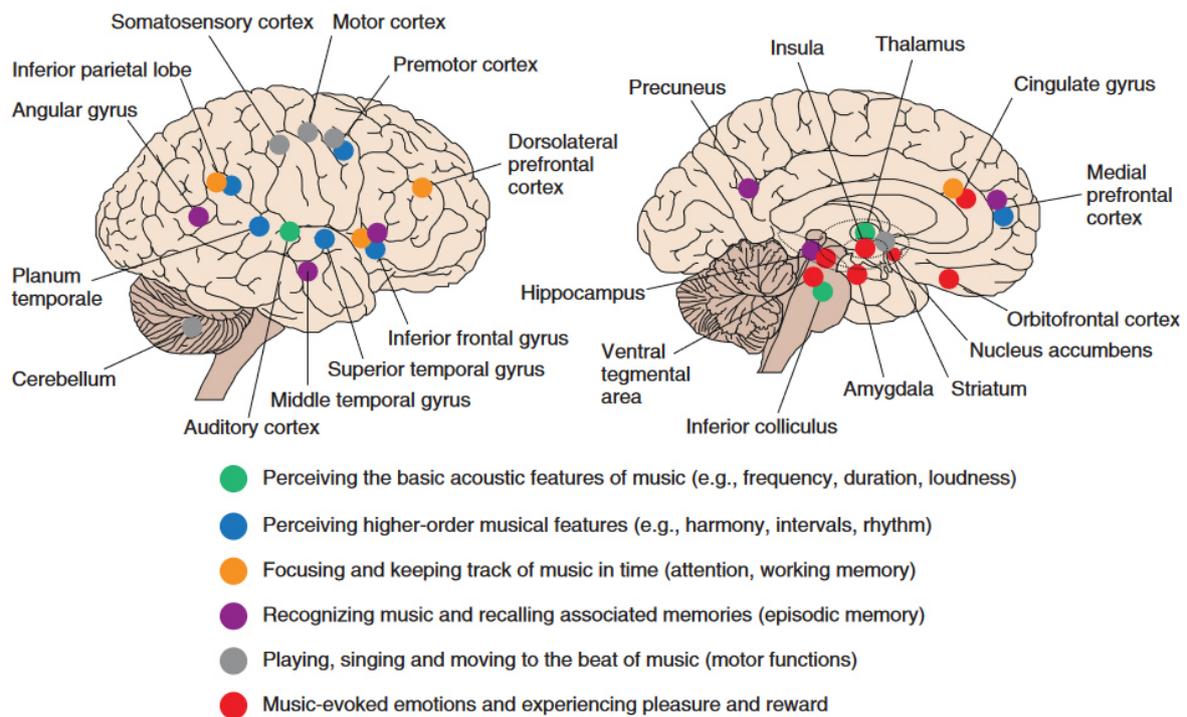


Figure 14: Cortical areas dedicated to the musical processing of information

(Reproduced with permission via Copyright Clearance Center: Särkämö et al., 2013; Copyright © 2013 by John Wiley and Sons.)

The processing of musical emotions involves the reward system, mostly when positive emotional arousal is triggered by the predicted resolution of musical uncertainty (Gold *et al.*, 2019; Salimpoor *et al.*, 2009; Zatorre and Salimpoor, 2013) and an extensive network which complexity may be due to the variability of internal (relative to felt emotions) and external (relative to music construction) emotions that are evoked by music. It involves, anterior Hipp, amygdala, auditory cortex, ventral (with accumbens nucleus, NAcc) and dorsal striatum, medial and lateral OFC, medial PFC, ACC, anterior insula, pre-supplementary motor area, rostral cingulate zone, medio-dorsal thalamus, cerebellum, motor areas... (Janata, 2015; Koelsch, 2018; Vuilleumier and Trost, 2015; Wang and Agius, 2018; Zatorre and Salimpoor, 2013). Some authors have proposed to distinguish the involvement of these areas into different processes. Stefan Koelsch (2018; 2015) explains music-evoked emotion in its quartet theory of emotion, in which he distinguished four networks interacting with each other (Figure 15, A). These networks, whose activity are presented to be regulated by amygdala, basal ganglia, striatum, insula, and/or cingulate cortex, were the following: brainstem-centered (arousing, effects, dissonance/consonance, fear), a diencephalon-centered (pain/pleasure), a hippocampus-centered (attachment-related affects), and an orbitofrontal-centered (social affects) affect systems. Another model, from Patrik Vuilleumier and Wiebke Trost (2015), proposed that the different categories of emotions evoked by music involved at different degrees areas that are activated for arousal, positive, and negative valence (Figure 15B). For example, wonder, a highly emotional and positive feeling, is associated to activation in the

ventral striatum and the hippocampus. The hippocampal activation is shared between all positive emotion. However, for low-arousing positive emotion, hippocampal activation is complemented with other regions involved in memory such as the PHC and the vmPFC, with a decrease activation in the ventral striatum.

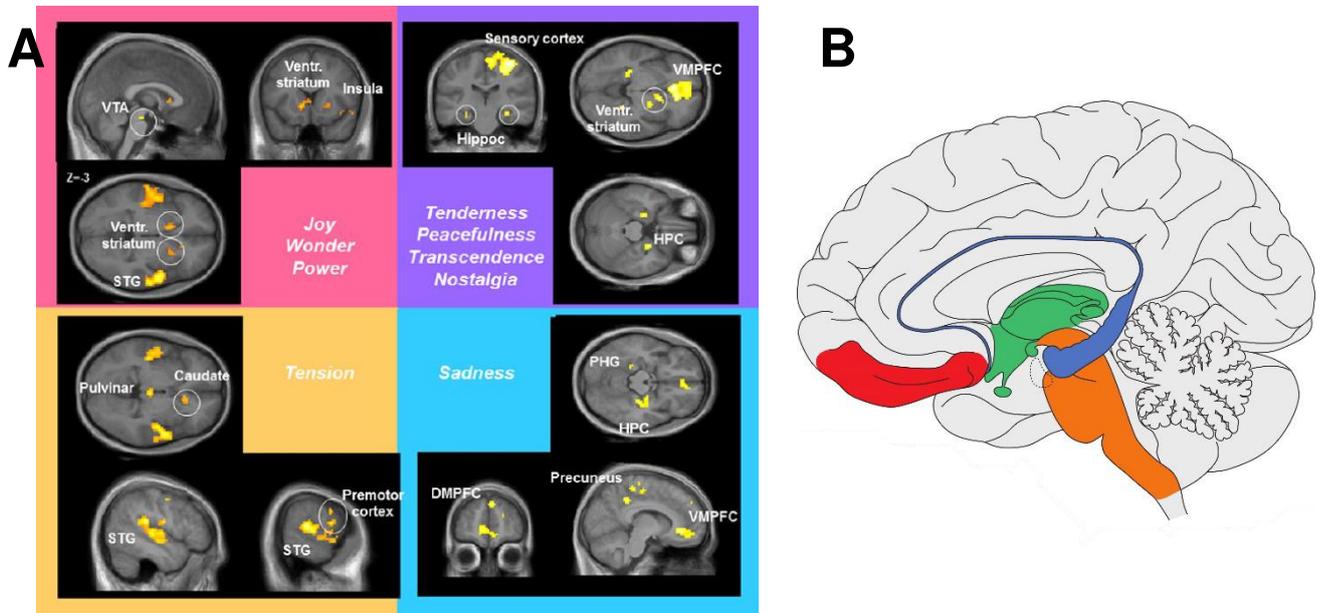


Figure 15: Cortical areas dedicated to the emotional processing of musical information

A. Pattern of brain activation for four types of emotion: positive high- and low-arousing emotions are in the two upper quadrants (pink and violet, respectively), and negative high- and low-arousing are in the two lower quadrants (orange and blue, respectively). B. The four affect system defined by Stefan Koelsch and collaborators: brainstem-centred (orange), diencephalon-centred (green), hippocampus-centred (blue), and orbitofrontal-centred (red). Activity in all these systems is changed by music listening. (Reproduced with permission: A. Vuilleumier and Trost, 2015; Copyright © 2015 by John Wiley and Sons & B. Koelsch et al., 2015; Copyright © 2015 by Elsevier)

3.2.3 Music and memory

3.2.3.1 Memory of music

The memory of music is fast: only two to three notes are sufficient to tell if the upcoming music is familiar, and two to three more notes to be able to tell its title (Schulkind, 2009). Music memory (AM) is also resistant, it is somewhat preserved in patients suffering for dementia such as Alzheimer disease (Cuddy, 2018, 2018; Platel and Groussard, 2010). However compared to other senses, in general population, such as vision (by comparison to images of object, faces, abstract art pieces, and degraded visual images of scenes), music recognition performance seems to be poorer (Cohen et al., 2009, 2011), or similar (Schulkind, 2009).

To unravel musical properties that contribute to music memory, research has mainly used the recognition Yes/No paradigm (Peynircioğlu et al., 2018) and tested the recognition of unfamiliar musical clips that have been changed on one dimension (*i.e.*, timber, key, tempo) between encoding and recall. Using this paradigm, Glenn Schellenberg and collaborators (2015; 2014)

have shown that relational structures of key and tempo were encoded between one day to one week of delay, as one week after the first exposure the change of key or tempo do not alter the recognition of the music, whereas it does after one day or 10 min. On the contrary, they have shown that specific timbre seems to be encoded directly, as a change of timbre was shown to alter the memory whatever the delay. This result was further supported by Kathrin Lange and Daniela Czernochowski (2013). In other words, a piece of music can be recognized even with tempo or key being changed, but not timbre. Music recognition indeed relies on perceptual features, as well as relational representations or structural relations (Peynircioğlu *et al.*, 2018). Isabelle Peretz and Robert J. Zatorre went further, proposing that the recognition of a musical clip relies on both abstract and absolute musical features, that are part of a musical semantic system called the *music lexicon* (Peretz and Coltheart, 2003; Peretz and Zatorre, 2005; Platel and Groussard, 2010; Cuddy, 2018). This system of musical knowledges (*i.e.*, familiar tunes, structural rules, timbres) that are implicitly constructed through mere exposure to music is thought to be different to, while sharing links with, the verbal semantic system. (Peynircioğlu *et al.*, 2018). Information stored in the musical lexicon allows for the detection of structural violation, the recognition and feeling of familiarity of previously heard musical tunes, and the recall of associated lyrics and AM (for a review, see Cuddy, 2018). Explicit knowledge about music are also encoded in the musical lexicon, explaining that musicians might have a better recognition memory for music than non-musicians (Cohen *et al.*, 2011; Talamini *et al.*, 2017; Weiss *et al.*, 2012).

As seen in the section 3.2.2.1, the links that emotion and music shared are strong and diversified. At the time of writing, 12 studies have investigated the effect of emotion, whether they are felt or recognized, on music recognition (Alonso *et al.*, 2015; Altenmüller *et al.*, 2014; Aubé *et al.*, 2013; Eschrich, 2005; Eschrich *et al.*, 2008; Ferreri and Rodriguez-Fornells, 2017; Ferreri *et al.*, 2021; Narme *et al.*, 2016; Nineuil *et al.*, 2020; Samson *et al.*, 2009; Stalinski and Schellenberg, 2013; Vieillard and Gilet, 2013). When the emotion felt by the listener is taken into account, recognition is favored by music liking (Stalinski and Schellenberg, 2013) - an effect that may be related to the reward value of the music (Ferreri and Rodriguez-Fornells, 2017; Ferreri *et al.*, 2021) -, emotional intensity (Aubé *et al.*, 2013 did show a correlation; Eschrich *et al.*, 2008 did show a tendency), and arousal (Alonso *et al.*, 2015; Ferreri and Rodriguez-Fornells, 2017; Ferreri *et al.*, 2021). Results are sparser when emotion is not evaluated by the participants but inferred from the material. Regarding arousal (*i.e.*, music intended to be happy *versus* music intended to be peaceful) some study did find an effect (Alonso *et al.*, 2015; Narme *et al.*, 2016), whereas others did not (Aubé *et al.*, 2013; Nineuil *et al.*, 2020). Concerning inferred valence, performance is sometimes better with positive compared to neutral (Eschrich *et al.*, 2008) and negative stimuli on the average, however this

effect seem to depend on retention duration, arousal level (Alonso *et al.*, 2015; Nineuil *et al.*, 2020), and the number of events (for example notes or chords) heard (Aubé *et al.*, 2013). To note, among cited studies, only few compared music to other stimuli (Cohen *et al.*, 2009, 2011; Schulkind, 2009), and none tested the effect of emotion conveyed by music versus other sensory modalities on recognition performance.

3.2.3.2 Memories evoked by music

Music is a particularly effective cue for AM. As a contextual cue, music is able to help patients suffering of dementia to recall memories while being exposed to a chosen song (El Haj *et al.*, 2012). In the general population, music is a powerful autobiographical memory cue, able to evoke AM at a high rate. In the laboratory, between 30% and 50% of presented excerpts were associated to an AM (Janata, 2009a; Janata *et al.*, 2007; Krumhansl, 2017). During everyday life, at least one music-evoked memory could be reported a day (Jakubowski and Ghosh, 2021).

Music-evoked remembered events come mostly from the adolescence (13-19 years) and young adulthood (20-29 years) periods, whether it was remembered with the musical sound stimulus itself (Krumhansl, 2017; Krumhansl and Zupnick, 2013) or only its title (Jakubowski *et al.*, 2020). To note, these results have been calculated with the release date of the song, and not the age participants gave to their memory, as it is done in other studies, which renders the comparison difficult with other sensory modalities (Chu and Downes, 2000a; Willander and Larsson, 2006, 2007; Willander *et al.*, 2015). Moreover, this bump in memory report may be explained by other factors, such as the amount of listening per day (Jakubowski and Ghosh, 2021; Krumhansl, 2017), music familiarity (Jakubowski *et al.*, 2020) and/or preference (Krumhansl, 2017). All these factors were indeed shown to be evaluated higher in the adolescence (13-19 years) and young adulthood (20-29 years) periods than in other periods of life.

Music evoked most frequently social situations, involving family and/or friends, and periods of life (Jakubowski and Ghosh, 2021; Janata *et al.*, 2007). Music seems to induce more positively than negatively valenced AM, with more memories rated as being happy or nostalgic (Jakubowski and Ghosh, 2021; Janata, 2009b; Janata *et al.*, 2007). The emotion of the retrieval cue was shown to influence the phenomenology and content of AM. The emotional intensity, then the positive valence, and then arousal, were shown to be potent dimensions that explain autobiographical salience of musical pieces (Salakka *et al.*, 2021). Moreover, emotional music evoked more detailed AM than less emotional ones (Janata *et al.*, 2007). However, whether this effect was due to an easier access to memory or through arousal or emotion conveyed by music itself (El Haj *et al.*, 2012; Schulkind, 2009; Sheldon *et al.*, 2020) is not known yet. Finally,

song's familiarity was shown to be linked with temporal specificity and valence of the evoked memory, with increasing familiarity being associated with memories that are more specific and positive (Ford *et al.*, 2016).

Only few studies compared music-evoked memories to memories evoked by other types of stimuli. Musical clips were shown to evoke memories that contain more perceptual details and a greater ratio of episodic information over total details than do memories evoked by faces (Belfi *et al.*, 2016), although faces seem to be associated with a stronger feeling of being brought back in time (Cady *et al.*, 2008). Similarly, music-evoked memories contain also more perceptual and social details, and are more vivid and more significant, more pleasant and emotionally intense, and give a better sense of reliving when compared to episodic memories cued by a TV show (Jakubowski *et al.*, 2021). However, music-evoked memories seem to be similar than to those evoked by words or lyrics (Cady *et al.*, 2008; Jakubowski *et al.*, 2020).

3.2.3.3 Neural bases of music memory

Correct recognition of a musical piece, whether achieved with a recollection or a familiarity based process, has been shown to rely on the right Hipp, the bilateral lateral temporal regions, the left inferior frontal gyrus, the left precuneus (Watanabe *et al.*, 2008), the cingular cortex whatever the cultural background (Demorest *et al.*, 2010), and the cerebellum (Altenmüller *et al.*, 2014). Regions that are involved in recognition differ as a function of the song's familiarity. Using recognition on familiar and non-familiar pieces of music, Hervé Platel and collaborators (2003) have indeed shown that the recollection of familiar music was right lateralized and involved precuneus and superior frontal gyrus, whereas recognition of unfamiliar music was involved bilateral regions including superior and middle frontal gyri and medial frontal cortex. These regions were further described in research in which the feeling of familiarity was investigated as a long term recognition task (Peretz *et al.*, 2009; Plailly *et al.*, 2007)?

The recognition of a musical piece can be done on a purely perceptual basis, that is when there is no associated retrieval of a label, lyrics, or AM. This process is thought to rely on a restricted musical semantic memory system (Cuddy, 2018; Platel and Groussard, 2010). It is thought to rely on a familiarity-based process that is specific to musical material for several reasons. First, music recognition memory is spared in the case of bilateral hippocampal lesion that impaired visual memory and AM (Esfahani-Bayerl *et al.*, 2019). Second, music recognition memory is also spared in Alzheimer's disease while patients have more problems recognizing verbal stimuli (Cuddy, 2018; Groussard *et al.*, 2019). Finally, a dissociation between familiarity-based and recollection-based processes had been shown by selectively impairing recollection-based recognition using transcranial direct current stimulation on the right parietal cortex

(Schaal *et al.*, 2015). However, this is the only study that investigated brain areas involved in familiarity versus recollection of music recognition memory.

The recall of music-evoked AM activates the same regions as other non-musical AM: PCC, lateral temporal lobes, parahippocampal gyrus, mPFC (Ford *et al.*, 2011). Specifically, the association between the musical cue and the memory was presented to rely on the mPFC (Janata, 2009a). Its involvement is thought to be specific to music-evoked AM as mPFC lesion does not have an effect on face-evoked AM, while it hampers music-evoked AM (Belfi *et al.*, 2018). Medial temporal lobe was not found in all studies investigating music-evoked AM. For example Jaclyn Hennessey Ford and collaborators (Ford *et al.*, 2011) did find a MTL activation, whereas Petr Janata (2009a) did not. This discrepancy may be related to increased specificity, as the study by Jaclyn Hennessey Ford and collaborators measured this factor, while Petr Janata did not. The discrepancy may also be related to the retrieval intentionality or effort, since participants in Petr Janata's study were not asked to retrieve a memory for each musical excerpt, while music-evoked AM were intentionally retrieved in the Jaclyn Hennessey Ford and collaborators' study.

3.3 The case of odor

3.3.1 *Importance of odors in human everyday life*

Odor, smell, fragrance, scent, aroma, all these words refer to olfactory perception, which has been falsely defined as being poor, crude and useless (McGann, 2017; Saive *et al.*, 2014b). On the contrary to these beliefs, odors are essential in our everyday life. It has been involved in three main functions: ingestive behavior, environmental hazards avoidance, and social communication (Croy *et al.*, 2014; Stevenson, 2010). Odors are a crucial compound of the food perception during eating, called the food flavor (Goldberg *et al.*, 2018). Indeed, the fact that, following mastication, molecules released by food odors are conveyed by expiration into the nose - forming the aroma of food - is a cause of conflation between taste and olfactory properties of food. While taste refers to only few sensations (salty, sweet, sour, bitter, and umami), humans are able to smell many more odorants (Bushdid *et al.*, 2014), making our food experience rich and diversified. This diversity allows to enjoy feeding experience, but is also a way to regulate our appetite (Stevenson, 2010). Regarding environmental hazards, odors can become warning cue to avoid eating spoiled food, or to be alerted by the smell put in town gas for instance. As a social communication mean, it is involved in reproduction with for example human mate selection and emotion communication (Ferdenzi *et al.*, 2020). Moreover, odors are also widely used as a healing or well-being agent (aromatherapy and aromachology) capable of changing mood, physiology and behavior (Angelucci *et al.*, 2014; Herz, 2009) and for aesthetic purposes when odors are used as perfume (Majid, 2021a).

Despite the wide panel of functions of the sense of smell, it is usually when people lose it (anosmia: total loss; hyposmia: reduced detection of odors) that they become aware of its importance in their daily life (Toller, 1999). When smell is lost by accident or following a disease, patients indeed suffer from psychological and social consequences: feeling of vulnerability (leading to environmental anxiety), worse relationship with friends and romantic partner, weight disturbances and eating disorders, and sometimes moderate depression (Croy *et al.*, 2014; Hummel *et al.*, 2017; Manesse *et al.*, 2021; Toller, 1999). Suffering endured by these patients were mainly seen as aesthetic and not important by the medical profession until recently (Toller, 1999). The COVID-19 crisis had made people more aware of negative consequences of smell loss and its importance in everyday life and in industry (Cameron *et al.*, 2021).

3.3.2 *Odor perception, emotion and cognition*

3.3.2.1 *Characteristics*

Contrary to false beliefs, humans do not have poor olfactory abilities (for reviews: McGann, 2017; Yeshurun and Sobel, 2010). Humans have high ability to tell that an odor is present in

the environment (detection), for instance being able to detect very low thresholds of odorant, the ethyl mercaptan, such as three drops of odorant in an Olympic-size swimming pool. Humans have almost perfect . Odor discriminating abilities, the ability to detect a difference between two perceived odors, are also very acute: molecules as closed as enantiomers (mirrors molecules) or differing only by the number of carbons in their carbonated chains or by their functional group, can be perceived differently. Going further, some authors proposed that humans should be able to discriminate one trillion olfactory stimuli (Bushdid *et al.*, 2014), however this number has been controversial (Meister, 2015), and evidence shows more surely that it is possible to discriminate more than 10,000 odors. These detection and discrimination abilities are highly variable between individuals (Mantel *et al.*, 2019). This particularity of olfaction in comparison to other sensory systems, arises from varied human olfactory receptor expression (Keller *et al.*, 2007) making humans with a normal sense of smell (normosmic) unable to detect certain specific odors (*specific anosmia*), or smelling it differently from one people to another. As an example, androstenone can be perceived as a floral odor, a sweat and urinous odor, or be the subject of specific anosmia. Androstenone can however be still detected by participants being anosmic to that odor by the involvement of the trigeminal system (Boyle *et al.*, 2006). Trigeminal system is closely related to the olfactory system (Hummel *et al.*, 2009), and allows for somatosensory perception that comes often when smelling an odor. It is involved for example in the sensation of freshness, or irritation that comes when smelling mint or mustard (Gerhold and Bautista, 2009). The smelling of an odor is thus the unitary perception of trigeminal and olfactory sensations.

Odor perception is holistic and associative. Odors are indeed perceived as a whole, as it is the case for faces. However, odor perception is thought to be even more synthetic than face perception (Behrmann *et al.*, 2014). As it is possible to perceptually divide every element present on a face, it is highly difficult to describe odor components (Olofsson and Gottfried, 2015; Wilson and Stevenson, 2003). Both trained participants and experts are unable to perceive more than three elements in an odor mixture (Livermore and Laing, 1996), maybe because the perception of an odor mixture differs from the separate perception of its sub-elements (Gottfried, 2010). The odor perception is further enriched by its inherent association with memory traces that have been previously associated to it, and with the perceiver's mental and internal states (Wilson and Stevenson, 2003). An odor will be smelled differently when hungry or starved, or if it was associated with food that has provoked an indigestion (Cabanac, 1971). This makes the odor perception highly personal, and may explain another facet of the high inter-individual variability associated to odor perception (Mantel *et al.*, 2019).

Humans demonstrate little ability to *identify* odors: less than 50% of our daily life odors can effectively being named (Engen, 1987; Majid, 2021a; Olofsson and Gottfried, 2015; Yeshurun

and Sobel, 2010). Similarly, it is also difficult to verbally categorize and communicate about odors (Jraissati and Deroy, 2021; Majid and Burenhult, 2014). Retrieving semantic information about an odor without the odor name appears to be difficult (Stevenson and Mahmut, 2013). This low ability can be explained by at least three hypotheses (Majid, 2021a, 2021b; Olofsson and Gottfried, 2015; Olofsson and Pierzchajlo, 2021). The first one is defended by Asifa Majid and her collaborators. She proposed that the poor ability to name odors is due to western cultures marginalizing the sense of smell in comparison with other ones, such as Jahai or Semaq Beri people in Malaysia. According to Asifa Majid, odor marginalization leads to poor attention and vocabulary is dedicated to it (Majid, 2021a, 2021b). In the second hypotheses, Jonas Olofsson and colleagues argued that this low ability to name an odor is more surely due to neural pathway that only allows for a weakly elaborated link between olfactory and semantic brain areas (Olofsson and Gottfried, 2015). Other explanations on this poor ability to name and communicate about odors may be our limited ability to mentally imagine a smell, that appears to be impossible while being feasible (Arshamian and Larsson, 2014; Royet *et al.*, 2013).

The last and maybe the most important particularity of olfactory perception is its links with emotional processing, emotional response being a whole part of olfactory phenomenological experience (Kontaris *et al.*, 2020; Richardson and Zucco, 1989; Yeshurun and Sobel, 2010). Among emotion dimensions, pleasantness (valence) is particularly important. Odors can be described and categorized according to their valence (Jraissati and Deroy, 2021), which is the first dimension by which an odor is perceived (Khan *et al.*, 2007; Zarzo, 2008). This strong link between olfaction and emotion has led Yeshurun and colleagues to propose that odors are exclusively coded in terms of pleasantness (Yeshurun and Sobel, 2010). All odors have thus an intrinsic emotional value (Kontaris *et al.*, 2020; Royet *et al.*, 2003). There is a debate in the literature trying to tell if this link between odor and emotion is innate or acquired. Olfactory preferences are mainly shared between species (Mandairon *et al.*, 2009), and some links between odor physicochemical properties and pleasantness had been unraveled (Khan *et al.*, 2007), suggesting that odor preferences are, at least for a part, determined at birth. However, most studies suggests that emotional responses to odors are, for the most part, acquired. Odors are indeed extremely sensitive to evaluative conditioning, the fact that a neutral odor will be strongly associated to the emotion lived the first time this odor was encountered (Herz, 2005; Herz *et al.*, 2004a). For example, odors are more prone than sounds to be associated with the emotion of a visual stimulus (Pützer *et al.*, 2019). Odor perception is also influenced by mood, and the opposite is also true: pleasant and unpleasant odor can respectively induce positive and negative mood (Kontaris *et al.*, 2020).

3.3.2.2 Neural basis of odor processing

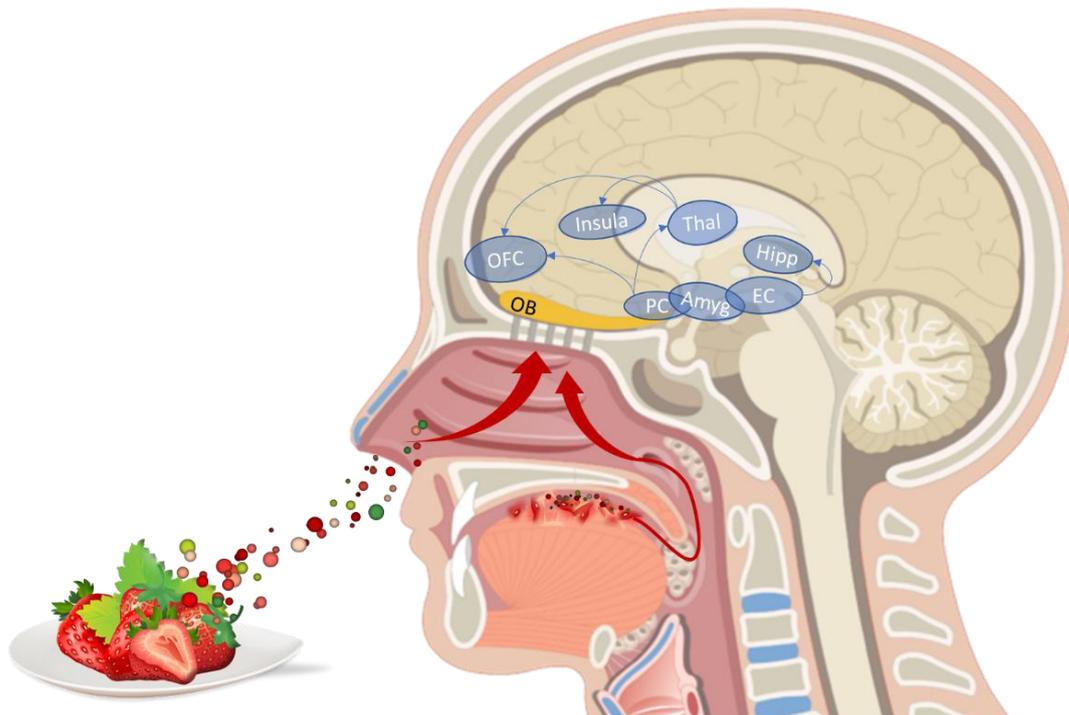


Figure 16: Schematic view of the human olfactory system.

The olfactory information enters the olfactory system either through the nose (orthonasal pathway), or through the mouth (retro-nasal pathway). The information is first processed in the olfactory bulb, and then in the primary (PC, Amyg, EC) and in the secondary (Insula, OFC) olfactory areas. OB: olfactory Bulb; Amyg: amygdala; EC: entorhinal cortex; Hipp: hippocampus; OFC: orbitofrontal cortex; PC: piriform cortex; Thal: thalamus. (Rights: adapted from Royet *et al.*, 2015)

Odorants are volatile molecules that can bind to and stimulate olfactory receptors located in the olfactory epithelium, located at the top of the nasal cavities. To reach the olfactory epithelium, odors are conveyed by airflow that can take two routes: orthonasal (a sniff or an inspiration convey the airflow through the nose) and/or retronasal (an expiration conveyed the airflow from the mouth through the pharynx to the nose) pathways (Hummel *et al.*, 2017). Odorants are then trapped in the olfactory mucosa, where each odorant can bind to different types of olfactory receptors, each type of olfactory receptor being able to bind to many odorants with different affinity (Reisert and Restrepo, 2009). As humans express between 300 and 400 olfactory receptors, odorants are represented through highly specific spatio-temporal maps (Hummel *et al.*, 2017; Lane *et al.*, 2020). This pattern of activation is transmitted to the olfactory bulb (OB), the first intracranial relay in the olfactory system. The signal is then sent to the primary olfactory cortex by the lateral olfactory tract, and result in the perception of an odor, the interpretation of the odorant by the brain (Wilson and Sullivan, 2011). Importantly and contrary to other sensory modalities, the information coming from OB does not go through a thalamic relay before reaching the primary sensory cortex (Baudry, 2020; Gottfried, 2010; Lane

et al., 2020). Another particularity is that the primary olfactory cortex is an assembly of anatomically and functionally diverse brain areas that all receive direct inputs from the OB: anterior olfactory nucleus (**AON**), olfactory tubercle (**OT**), piriform cortex (**PC**), cortical nucleus of the amygdala, and entorhinal cortex (**EC**). These structures send then information to secondary regions that include lateral and basolateral amygdala, agranular insula, hypothalamus, PRC, Hipp, striatum and OFC (Gottfried, 2010; Hummel *et al.*, 2017; Kontaris *et al.*, 2020; Lane *et al.*, 2020; Wilson and Sullivan, 2011; Yeshurun and Sobel, 2010).

Primary olfactory cortex brain regions have different functions⁹ (2020; Lane *et al.*, 2020; Zhou *et al.*, 2019). AON is involved in odor object representation and location. Its links with OFC and Hipp are thought to be respectively responsible of its involvement in odor recognition and episodic memory integration. OT is mostly known by its links with the reward system and is thus thought to take part in the attribution of a reward value to an odor object and motivation processes linked to odor perception. OT also sends projections to motor regions, paracingular cortex, left frontal pole, and left fusiform gyrus and is thought to support odor-related attention and social processes.

PC has been extensively studied being often referred to as the primary olfactory structure that is responsible to represent the odor object. However, this region is also considered to be an associative area rather than a pure primary cortex. It is indeed involved in many complex cognitive abilities such as working memory, imagination, learning and attention related to odor processing (for reviews, see: Gottfried, 2010; Lane *et al.*, 2020), and also in sniffing and emotional breathing related activity by its links with amygdala (Masaoka *et al.*, 2014). In human, PC can be divided into two regions, the frontal and the temporal piriform cortices, two areas that may correspond to anterior and posterior piriform cortices in rodents (Wilson and Sullivan, 2011). According to Guangyu Zhou and collaborators (2019), the frontal piriform cortex would be involved in motor planning in response to odorant perception as it sends projections to motor planning areas (putamen, primary motor cortex and caudal node) and left supra marginal gyrus (tool grasping). Frontal piriform cortex also represent perceived or imagine odor pleasantness (Lane *et al.*, 2020). PC have indeed reciprocal connections with regions involved in emotional processing, such as OFC and basolateral amygdala. Temporal PC, through its projection to brainstem raphe magnus, posterior insula and core language network, would be responsible to olfactory verbalization and olfactory modulation of breathing.

Anterior cortical nucleus of the amygdala is linked to innate aversive and attractive reaction to odor stimuli in rodent literature. This region, with OFC and PC, sends strong reciprocal

⁹ The reviews of Ioannis Kontaris and its collaborators (2020), of Gregory Lane and its collaborators (2020), and of Guangyu Zhou and collaborators (2019) guided the construction of the following part about primary olfactory cortex brain regions functions.

connections with basolateral amygdala and other well-known regions that are part of the *affectome* (*i.e.*, regions involved in emotional processes Becker *et al.*, 2019), providing more proof to the closed links between emotional and olfactory processing. It takes thus only three synapses to an odor signal to reach the basolateral amygdala, that is involved in the relationship between odor perception and emotion (valence and emotional intensity), and mood (Kontaris *et al.*, 2020; Royet *et al.*, 2003; Yeshurun and Sobel, 2010). Amygdala shows also high functional connection with areas that are involved in memory, including Hipp and EC (Arnold *et al.*, 2020).

3.3.3 Odor and memory

3.3.3.1 Memory of odors

Numerous studies evaluating memory of odors rely on recognition or association paradigms that have compared odors to images or word cues (Davis, 1975, 1977a; Engen, 1987; Herz and Cupchik, 1995a; Murphy *et al.*, 1991; Olsson *et al.*, 2009). These studies showed that odors were less recognized than concrete visual cues (Cornell Karnekull *et al.*, 2015), but not than abstract visual forms (Lawless, 1978; but see: Davis, 1977), and environmental sounds (Cornell Karnekull *et al.*, 2018). Odor memory is influenced by many factors. First, it deeply relies on breathing, recognition being better during nasal than mouth breathing (Arshamian *et al.*, 2018). Odor memory is also highly influenced by its conscious identification, as familiar and identified odor are better recognized than unidentified ones (Cessna and Frank, 2013; Cornell Karnekull *et al.*, 2015). It has notably been shown that memory for identified odors is similar to memory for words, but less good than unidentified odor (Olsson *et al.*, 2009). Finally, and importantly, odor memory is influenced by emotion. Valence influenced odor recognition memory such as bad odors are better recognized than neutral or positive odors (Larsson *et al.*, 2009). Anne-Lise Saive and collaborators (2014a) did not replicate these results but showed that both positive and negative odors are more potent cues than neutral ones. These discrepancies may come from another emotional dimension that was lacking in these studies: emotional arousal (or intensity). Emotional intensity might have helped memory processing instead of valence.

Opposite to their somewhat limited ability to be recognized, odors were shown to be potent memory cues. For example, reinstating the encoding olfactory context at recall significantly improve participants' performance (Chu and Downes, 2000b). They are as potent cue as visual, tactile and verbal versions of the olfactory objects to recall the associated item (Davis, 1977; Herz, 1998; Herz and Cupchik, 1995; but see: Davis, 1975; see also Engen, 1987 for a review). Rachel Herz and colleagues have also shown that the memory of pictures associated with odors were more emotionally loaded than the memory of pictures associated with musical,

verbal or tactile cues. This emotion effect seemed to have an influence on confidence but not performance (Herz, 1998; Herz and Cupchik, 1995a).

3.3.3.2 Memories evoked by odors

A common report about odor memory is the ability of odor to evoke singular memories of personal events, which is often referred to as *The Proust phenomenon* (Chu and Downes, 2000b, 2002). Marcel Proust, in *Swann's way* (1913), described with many details its observation and feelings at the onset, search and retrieval of a memory that has been induced by eating a madeleine just dipped into its tea. However, as this episode referred to a multisensory cue and not a solely olfactory cue, and as the phenomenon described by Proust is far more complex than most of odor-evoked memories (for review, see: Jellinek, 2004) it will be the only reference to the Proust phenomenon in this manuscript.

The particularity of odor-evoked memories towards memories evoked by other sensory cues (for reviews, Annett, 1996a; Hackländer et al., 2018; Herz and Engen, 1996a; Larsson et al., 2014; Saive et al., 2014b) has been addressed comparing voluntary odor-evoked memories of participants' life events to those evoked by labels and/or images (de Bruijn and Bender, 2017; Chu and Downes, 2000a, 2002; Glachet and El Haj, 2021; Goddard et al., 2005; Herz and Schooler, 2002; Herz et al., 2004b; Hinton and Henley, 1993; Miles and Berntsen, 2011; Rubin et al., 1984; Willander and Larsson, 2006); sound and/or images (EL Haj *et al.*, 2018; Herz, 2004; Karlsson *et al.*, 2013; Knez *et al.*, 2017; Toffolo *et al.*, 2012; Willander *et al.*, 2015) and more recently the four other sensory modalities (Ernst *et al.*, 2021; Hutmacher, 2021). Among many phenomenological characteristics that have been addressed in these works, it has been consensually admitted that olfactory cues evoke less memories, and are older and more emotional than those evoked by other sensory modalities (Larsson et al., 2014).

Odors are shown to often cue less memories of life events than cues from other sensory modalities (Ernst *et al.*, 2021; Goddard *et al.*, 2005; Hinton and Henley, 1993; Willander and Larsson, 2007; but see: Ernst *et al.*, 2021, for taste cues; Miles and Berntsen, 2011; Willander and Larsson, 2006; Willander *et al.*, 2015). This result may be related with poorer recognition and identification performance for odors relatively to stimuli from other senses. For example, as Patricia Hinton and Tracy Henley (1993) discussed, participants, when discovering a cue, may do a mental blockage in their ability to name the odor, impairing thus the onset of the AM process.

Memories evoked by odors show a distinct distribution across life than other sensory modalities, with more memories recalling childhood period than more recent periods (Larsson et al., 2014). The frequency distribution of AM evoked by different cues across the lifespan highlights these differences. This distribution is characterized by three periods: childhood

amnesia, where no memories could be evoked; the early bump, where most of the memories come from; and finally, a recency period. The early bump is located in young adulthood for almost all tested sensory cues (verbal, auditory, visual and multisensory), although for olfactory cues, it was located before 10 years old, in the first decade of life (Chu and Downes, 2000a; Willander and Larsson, 2006, 2007; Willander *et al.*, 2015).

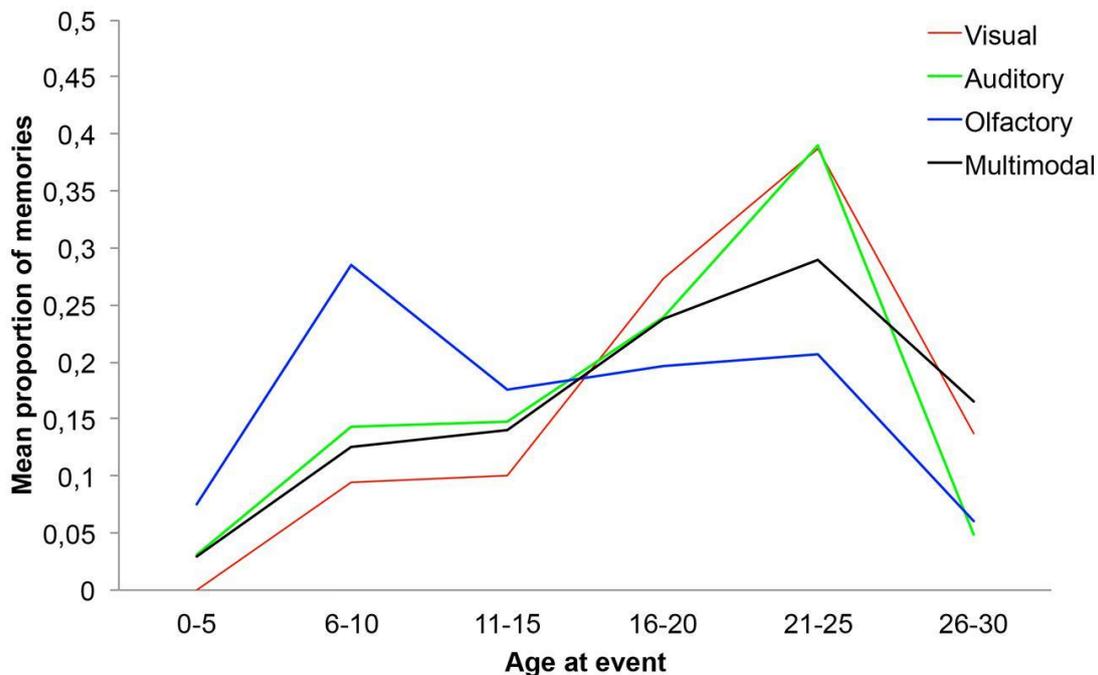


Figure 17: AM distribution across lifespan as a function of sensory modality

(Reproduced with permission: Willander *et al.*, 2015; CC BY 4.0)

Olfaction and emotion are closely linked, both phenomenologically with the emotional potency of the odor cue (Mandairon *et al.*, 2009), and by sharing common neural bases (Saive *et al.*, 2014b). Maybe due to their close links, odor-evoked memory are often presented as being particularly emotional. Whereas most studies report an emotional effect (Arshamian *et al.*, 2013; Chu and Downes, 2002; Herz, 2004; Herz and Schooler, 2002; Rubin *et al.*, 1984; Willander and Larsson, 2007) or trend (Miles and Berntsen, 2011), with odor-evoked AM being more emotionally loaded than AM evoked by other sensory cue, some do not (de Bruijn and Bender, 2017; Ernst *et al.*, 2021; Willander and Larsson, 2006; Willander *et al.*, 2015). However, there are several problems in these studies, and they could hardly be compared. First, emotion is investigated through different ratings: valence and emotional intensity, and/or a broader concept of emotionality. Next, as all studies did not investigate memories of unique events located in time and location but also repeated or long-lasting events, a potential confound can occur, as repeated and long-lasting events can each be composed of several and contradictory emotions. Also, most of the stimuli that are used, if not all, are positively valenced, restricting the diversity of memories that can be evoked. Finally, ratings could

concern emotion at encoding and/or recall, which is not always specified (de Bruijn and Bender, 2017). It thus difficult to have a clear idea about the emotional dimension that influences memory, and what are the underlying physiological and neurological processes. Studies done by Yuri Masaoka and her team can give clues. While not providing comparison to other senses, they show showed that odor-evoked AM are emotionally loaded such that they influenced physiological and phenomenological associated experience. In comparison to non-evoking odors, odors that evoke AM are associated to a deeper and slower breathing, and increased ratings of odor emotional arousal and pleasantness, memory arousal level, and familiarity feeling (Masaoka *et al.*, 2012; Watanabe *et al.*, 2018). Finally, odors are shown to be potent cues that can induce traumatic memories, notably by evoking flashback in post-traumatic stress disorders (Daniels and Vermetten, 2016; Vermetten and Bremner, 2003).

Many other characteristics had been investigated to characterize odor-evoked AM. They are often rehearsed with a stronger feeling of being brought back in time (Arshamian *et al.*, 2013; Herz, 2004; Herz and Schooler, 2002; Willander and Larsson, 2006, 2007), and only one study did not find this effect (Willander *et al.*, 2015). However, it is not clear whether odors cue influence memory vividness, some results showing that odor induced less (Goddard *et al.*, 2005; Willander *et al.*, 2015) or equally vivid (Herz, 2004; Herz and Schooler, 2002; Rubin *et al.*, 1984; Willander and Larsson, 2006, 2007) memories than verbal, auditory and/or visual cues. However, when childhood odors are compared to childhood pictures, memories were shown to be more vivid and of a higher quality (de Bruijn and Bender, 2017). The same is observed for memory specificity, odors cueing memory that are less or equally specific than verbal, auditory and/or visual cues (Ernst *et al.*, 2021; Goddard *et al.*, 2005; Herz, 2004; Herz and Schooler, 2002; Miles and Berntsen, 2011). These memories were also shown to be surprising when they come to consciousness as they were less rehearsed, talked about and thought of less often than memories evoked by other sensory modalities (Rubin *et al.*, 1984; Willander and Larsson, 2006). However, this surprise effect need to be further explored (Hackländer *et al.*, 2018). Among all presented studies that investigate odor-evoked AM, only few tempted to measure accuracy (Hackländer *et al.*, 2018). In fact, one ecological task (Chu and Downes, 2002) tempted to verify accuracy by comparing details that were first remembered with words with those evoked by other stimuli. However, this method did not measure accuracy but consistency of reports. Only one ecological-laboratory tasks investigated this showing the ability of odors to recall complex events, but did not compare odor-evoked memories to stimuli from other sensory modalities (Saive *et al.*, 2013, 2014a, 2015).

Several reasons had been raised to explain the specificities of odor-evoked AM (Hackländer *et al.*, 2018). Among them, it is proposed that the first odor-item association is strong (Yeshurun

et al., 2009). This can result in a stronger pro-active interference of odors than other sensory items, explaining for example the early childhood bump specific to olfactory cued memories. Also, semantic information affect as much laboratory (Cessna and Frank, 2013; Cornell Karnekull *et al.*, 2015; Olsson *et al.*, 2009) than more ecological odor-evoked memories. Indeed, when an odor is associated to a name, the phenomenological characteristics and the age distribution of odor-evoked AM are changed relative to when an odor is presented alone: the early bump is less pronounced, and the event emotional intensity and feeling of be brought back in time were evaluated as less important (Willander and Larsson, 2007). Thus, it is possible that an unidentifiable odor that comes out from the environment could be linked to life events (semantic processing differences hypothesis), and/or evoke specific memories (language by-pass hypothesis) differently than if the odor is identifiable.

3.3.3.3 Neural basis of odor memory

Odors, emotion and memory share common neural basis, that are thought to be involved in the distinctive nature of odor-evoked memories (for reviews: Daniels and Vermetten, 2016; Saive *et al.*, 2014a). Indeed, these regions are anatomically and functionally close together. Amygdala and EC, that are involved in emotion and memory, are both separated from the olfactory receptor neurons by three synapses. Moreover, Hipp, OT, nucleus accumbens, OFC subregions and hypothalamus, are all parts of one olfactory subnetwork (among three), that organized olfactory system during resting state (Arnold *et al.*, 2020). The few studies that compare the neural bases underlying the same memories evoked either by odors or by words or visual items sustain this hypothesis. They have shown that odor-evoked AM were associated with a greater activation of the amygdala, PHC, striatum, occipital gyrus (Herz, 2004), precuneus, insula, and temporal regions (Arshamian *et al.*, 2013), which are regions involved in recollection of AM, and in olfactory and emotional processing. It is also shown that the direct activation of the amygdala in an epileptic patient evoke olfactory hallucination followed by an odor-evoked reminiscence. This behavior was sustained by an increase of the connectivity between amygdala, insular cortex and temporal pole (Bartolomei *et al.*, 2017). Other authors, comparing non-evoking and an AM-evoking odors further explained the role of the dorsal OFC. Its involvement and its connectivity with the fusiform gyrus was related to a subjective feeling of remembering and being brought back in time (Masaoka *et al.*, 2021; Watanabe *et al.*, 2018).

Laboratory tasks are used to determine brain bases of olfactory memory accuracy. Participants having the best odor recognition performance involved a network composed of Hipp, caudate nucleus, ACC and medial temporal gyrus (Meunier *et al.*, 2014), Hipp and PHC being more activated in correct odor recognition than other memory response categories (Royet *et al.*,

2011). Hipp is indeed central in odor recognition (Levy, 2004) and odor-object association (Goodrich-Hunsaker *et al.*, 2009), as a lesion of Hipp impaired both. Hipp have also been shown to be more activated when the first cue that was associated to a picture was an odor *versus* a sound (Yeshurun *et al.*, 2009). In fact, the network involved in recognition seems to be a more global recognition network equally engaged whatever the cue, but with the addition of olfactory regions that are specific to olfactory retrieval (Lehn *et al.*, 2013). This network also partially overlap with regions that are involved in multimodal feeling of familiarity (Plailly *et al.*, 2007).

OBJECTIVES & HYPOTHESIS

4 A brief recall of the context

Memory can be distinguished into three serial stages: encoding, storage and retrieval. Memory can also be distinguished according to the duration of the storage, the behavior that is engaged at retrieval, and their neural basis. In this framework, the memory of personal past events is called EM. Episodic memory has been defined thanks to the research of Endel Tulving (1972) classified as a declarative and perceptual and non-semantic memory. EM recall involves the recollection of past and future events embedded in a spatio-temporal context. EM can also be defined in the framework of AM, in which it can be defined as a specific AM, but with the additional involvement of SM. In these two frameworks, EM and AM are associated with different kind of experimental paradigms. EM studies involve laboratory tasks that focus on the investigation of core EM processes through measures of performance and veracity, controlling every memory stage. AM studies involve ecological tasks that focus on the investigation of naturally encoded everyday events through the analyses of their retrieved content. Nowadays, more and more researchers tend toward more ecological approaches, and new paradigms has been set up. Laboratory-ecological approaches were set up to construct a bridge between EM and AM frameworks. This evolution in these paradigms is of great importance, given that laboratory and ecological tasks rely on weakly overlapping neural basis, laboratory tasks requiring more control processes and less involvement of the DMN than do ecological tasks. To study EM, it appeared important for us to choose an approach that allows for the exploration of EM complexity while controlling each of the memory stages. We thus choose to develop a paradigm based on previous research done in our laboratory using the WWW paradigm (Saive et al., 2013, 2014a, 2015) in VR.

The memory of personal events is reconstructive and labile. It has noteworthy been shown that sensory modality influences episodic AM. In this thesis, we compare the effect of three kinds of stimuli: face, music, and odor. These three stimulus types have not only a great importance in our daily life, but also particular links with the memory of personal events. **Faces** are central in human socialization and communication, because it is a mean to convey numerous information, such as gender, emotion, social category, *etc.* Recognition of unfamiliar faces is picture based, and each modification that occurs between encoding and retrieval impairs its recognition. As a cue to AM, unfamiliar faces seem to induce the retrieval of more details and an increased feeling of being brought back in time as do music stimuli. Regarding neural bases, faces involve occipital and temporal face areas. **Musical** pieces share close links with emotion and are used as a mean of communication that act differently from language. The memory of music is fast and resistant, and favored by music liking, emotional intensity, and arousal. As a cue, music evokes positively valanced events, that are more detailed when the music is emotional. Events recollected with a music cue are more perceptual and emotional than events

recollected with visual cues. Music-evoked brain activation includes an incredibly varied amount of brain regions, including motor areas and the reward system, particularity that may distinguish its processing from that of other sensory modalities. **Odors** are a warning cue, a social communication mean, or a way to color the world, by making the food enjoyable or improving wellbeing in the environment. They are emotional by essence and the first reaction to odor is often an approach/avoidance behavior. The recognition of odor stimuli is good, but not as good as the recognition of other sensory stimuli and it is influenced by their identification and emotion. When used as a cue to evoke AM, odors evoke memory that are more emotional, older, rarer and rehearsed with a stronger feeling of being brought back in time. Odors neural network is clearly distinguishable from that of other sensory modalities. First, the information coming from the olfactory receptors reaches the primary region by bypassing the usual thalamic relay. Second, the primary olfactory regions include limbic structures. Thus, the information coming from the receptors quickly reach the amygdala and the hippocampus, two core regions that are involves respectively on emotion and episodic AM.

5 Objectives and Hypothesis

Odors seem to have a particular link with the memory of one's personal past events, that may be driven by their emotional potency and its close neural links with limbic areas. We made the following hypothesis:

The strong links shared between olfactory and emotional processes promote the retrieval of episodic memories.

To test this hypothesis, we set up a protocol in VR that allow the comparison of visual, olfactory and auditory processing and adaptable to fMRI experiment. We adapted this protocol to understand how the sensory modality of the memory cue affect recognition and EM processes for neutral (**objective 1**) and emotional stimuli (**objective 2**). Finally, an fMRI study allowed to study the functional connectivity of the EM network as a function of performance and sensory modality (**objective 3**).

The first objective is investigated in **study 1**, in which the hypothesis was that **while odor recognition might be not as good as music or face recognition, episodic memory would be more efficiently evoked with odors than with music of faces cues**. In the here created paradigm, participants discovered three events for three days (one per day). Each event included associations of three stimuli from three different sensory modalities (*What*, face-music-odor association), placed in a particular location (*Where*) in a particular room of a house (*Which context*). Memory was investigated the fourth day, during which participant were cued by uni-sensory stimuli. They were required to tell whether they recognize the cue, and then to retrieve the room and the place in the room where it was presented at encoding. Finally, to

better understand the memory scores and potential modality effects, response times, physiological responses (breathing), stimuli sensory evaluations (pleasantness, emotional intensity, familiarity and complexity) and participants' characteristics (via a questionnaire), were assessed.

The second objective is investigated in **study 2** in which the hypothesis was that **while emotion (subjective valence and emotional intensity) improves both recognition and EM performance of cues from each sensory modality, this effect would be stronger for odor than music or face stimuli**. In this study, the neutral stimuli were changed for negative, neutral and positive stimuli, and a new protocol was created to limit the drawbacks identified from study 1. For example, the episodes were changed. Here, during an episode, the participants discovered the entire house presented in one out of three light conditions, defining three periods of the day (*Which*). The three rooms represented the location (*Where*), and in each room were presented three devices that contained each one sensory stimulus (either an odor, a music or a face). The fourth day the same task was proposed to the participants. Finally, to better understand the memory scores, response times, physiological responses (breathing, electrocardiogram, skin conductance), stimulus (pleasantness, emotional intensity, familiarity and wanting) and participants' characteristics (via questionnaires). To note, these physiological data are unfortunately not presented in this thesis.

The third objective is investigated in **study 3** in which the hypothesis was that **the particularity of odor stimuli as EM cues comes from a higher neural connectivity between primary sensory cortex, amygdala, and hippocampus than for cues from other sensory modalities**. To test it, positive and negative odors were compared to positive and negative music in the same task that was used in study 1. The recall occurred in fMRI, where participants performed a subjective memory task followed by an episodic memory task. Using a Dynamical Causal Modelling analysis on the functional data, we expected to see a higher connectivity between sensory and memory and emotion regions for odor cues than for music cues. We also expected to see a higher activation of the amygdala and the hippocampus in the olfactory than in the music condition during the memory access.

RESULTS

Study 1

Episodic memory for odors, music and faces– an ecological laboratory approach using virtual reality

Lucile Rey^{1-2*}, Clément Désoche³⁻⁴, Marc Thévenet¹⁻², Samuel Garcia¹⁻², Barbara Tillmann^{2**}, Jane Plailly^{1**}

¹ CMO team, Lyon Neuroscience Research Center, CNRS UMR 5292 - INSERM U1028 - University Lyon1, Lyon, France

² CAP team, Lyon Neuroscience Research Center, CNRS UMR 5292 - INSERM U1028 - University Lyon1, Lyon, France

³ IMPACT team, Lyon Neuroscience Research Center, CNRS UMR 5292 - INSERM U1028 - University Lyon1, Lyon, France

⁴Hospices Civils de Lyon, Mouvement et Handicap, Neuro-Immersion, Lyon, France

**The last two authors contributed equally to this work

*Correspondence:

Lucile Rey

Centre de Recherche en Neurosciences de Lyon, INSERM U1028 - CNRS 5292 - UCBL - UJM

Centre Hospitalier Le Vinatier, Bâtiment 462 - Neurocampus Michel Jovet

95 boulevard Pinel 69675 Bron Cedex FRANCE

mailto: lucile-rey@laposte.net

Running title: Episodic memory and sensory modalities

Abstract

Episodic memory allows for the conscious re-experience of unique personal events from the past embedded in their specific spatio-temporal context and can be defined as the memory of *what* happened, *where* and in *which context* or *occasion*. Episodic memory retrieval is reconstructive, and the cue's sensory modality can influence the access and the content of the retrieved memories. However, the effect of each sensory modality is not well known yet. This study investigated the effect of odor, music and face cues on the recognition memory and episodic memory retrieval in a laboratory-ecological task using a non-immersive virtual reality device presenting a three-room house. Participants freely and incidentally explored three unique rich episodes over three consecutive days. Episodes were constructed around the three dimensions that characterize episodic memory: *What* (odor, music, face), *Where* (location in a room), and in *Which context* (bedroom, living room, office). The retrieval was tested on the fourth day. Participants were told to 1) recognize encoded stimuli among distractors and, 2) to go to the room and select the box in which they encountered them at encoding. This paradigm and its analyses allowed us to study the episodic memory processes as a whole as well as step by step. Results revealed that odor, music and faces were associated to different memory patterns. Odors were less recognized than both music and faces, but only odor and faces were able to accurately evoke a complete episodic memory. For all three sensory modalities, correct recognition was accompanied with modulation of breathing. The investigation of the potential effect of participants' individual sensory evaluations of the cues on memory performance suggest the importance of emotional intensity, opening up perspectives for future investigation of episodic memory.

Keywords: Episodic memory; Recognition memory; Odors; Music; Faces; Virtual Reality; Breathing; Human.

1 Introduction

In a multisensory world, episodes of one's life are rich in sensations, perceptions and emotions. These episodes might be encoded, consolidated, stored in memory and re-evoked later on. Episodic memory (EM) is the memory that allows for the conscious re-experience of unique personal events from the past embedded in their specific spatio-temporal context. It can be defined by its content as the memory of *what* happened, *where* and in *which context* or *occasion* (Easton and Eacott, 2008a; Tulving, 1983). It has been opposed to semantic memory, the memory of facts and general knowledge (Tulving, 1972). For example, running into a friend in the street can trigger, after recognizing him/her as an old friend, the simple memory of his/her name (semantic memory) and/or the recollection of a shared life episode. EM retrieval is reconstructive, that is, memories are constructed out of the initially-encoded elements (Bartlett, 1932; Schacter and Addis, 2007). Several variables can influence the memories' content and the associated phenomenological features, such as the congruence between encoding and retrieval contexts in terms of the environment (Smith *et al.*, 1978; Tulving and Thomson, 1973) or the emotional status (Bower *et al.*, 1978), for instance. Memory retrieval is also influenced by the cue that triggers recollection. All details of the event could potentially serve as cues triggering memories of this event, and the recollected memories depend on the nature of the memory cue. The present study focused on one dimension of a potential cue for EM, notably the cue's sensory modality.

The cue's sensory modality has been shown to influence numerous aspects of recollected memories (Ernst *et al.*, 2021; Hutmacher, 2021). The most common example is the Proust phenomenon (Chu and Downes, 2000, 2002), suggesting the particularity of odor-evoked memories compared to memories evoked by other sensory cues (for reviews, Annett, 1996; Hackländer *et al.*, 2018; Herz and Engen, 1996; Larsson *et al.*, 2014; Saive *et al.*, 2014a). This hypothesis has been investigated by comparing voluntarily evoked memories of participants' life events that are triggered by odor cues to those evoked by cues in the form of images, sounds and labels (de Bruijn and Bender, 2017; Chu and Downes, 2002; Ernst *et al.*, 2021; Goddard *et al.*, 2005; Herz, 2004; Herz and Schooler, 2002; Karlsson *et al.*, 2013; Miles and Berntsen, 2011; Rubin *et al.*, 1984; Willander *et al.*, 2015). Among various phenomenological characteristics addressed in these studies, findings have congruently shown that while olfactory cues evoke less autobiographical memories, the evoked memories are more emotional, older, and less often rehearsed, thought and talked about than memories evoked by other sensory modality cues. The retrieval of odor-evoked episodic memories has also been rarely studied with memories encoded in the laboratory using association paradigms, where the ability to recall the items previously associated with a cue during explicit encoding was tested (Davis, 1977; Herz, 1998; Herz and Cupchik, 1995). In this experimental approach, EM

characteristics overlap with the specificities observed in autobiographical approaches. While odors are less recognized than visual cues when presented among distractors, they show similar ability to initiate the retrieval of previously associated items (paintings, numbers, pictures) than do visual, tactile, musical and verbal cues (Davis, 1977; Herz, 1998). The odor-evoked memories are also significantly more emotionally loaded than the memories cued by the other modalities (Herz, 1998; Herz and Cupchik, 1995). However, the very limited number of studies urges to be cautious about conclusions and encourages to continue to explore this field of research.

Odors are not the only potent memory cue, other stimuli with strong ecological validity, such as music and face cues, also demonstrate specificities for evoking EM. Music-evoked autobiographical memories are reported to be consistently associated with strong emotions (Janata *et al.*, 2007; Schulkind *et al.*, 1999). They are also suggested to be specifically highly vivid. Music-evoked memories contain more perceptual details and a higher ratio of episodic experiencing information over total details than do face stimuli (Belfi *et al.*, 2016). Music-evoked memories contain also more perceptual and social details, and are more vivid, more significant, more pleasant and emotionally intense, and gave a better sense of reliving when compared to episodic memories cued by a TV show (Jakubowski *et al.*, 2021). Faces are associated with more recollection (*vs.* feeling of familiarity) than are voices and both episodic and semantic information about an individual are more likely to be retrieved following face recognition rather than voice recognition, even when faces are blurred in order to limit recognition (for an overview, see Brédart and Barsics, 2012). Although these findings suggest that EM might be influenced by the sensory modality of the retrieval cue for odors, music and faces, the evidence for these cue effect differences between sensory modalities is still missing. Aiming to bridge this gap, our present study aimed to challenge odor-evoked EM specificities by directly comparing EM evoked by odors with the same memories evoked by other potent memory cues, notably music and faces.

Two main approaches have guided EM investigations (McDermott *et al.*, 2009). Naturalistic approaches focus on autobiographical memories investigating real-life event memories evoked by memory cues or interviews. Laboratory approaches investigate memories created in the laboratory, with recognition memory, Remember/Know, or paired associate paradigms. Each of these two approaches have strengths and drawbacks. While naturalistic approaches allow investigating EM in a more ecological way and including several dimensions, they lack experimental control on non-measured variables and accuracy measures. In contrast, laboratory approaches allow controlling all memory stages, but the memories are usually limited to a one-dimensional item. These differences with their potential effects on the specific features of EM reveal the necessity to develop more integrated investigation approaches,

notably by adding ecological validity to a controlled laboratory experiment. For this goal, we have set up new experimental protocols inspired by studies investigating episodic-like memory in animals (Allerborn *et al.*, 2016; Clayton *et al.*, 2001; Veyrac *et al.*, 2015) and based on content-based approaches of EM (Easton and Eacott, 2008b; Tulving, 1972, 1983). In our first developed protocol, participants freely explored an olfactory environment and were later on asked to remember olfactory episodes constructed around three questions: “*What* was the event”, “*Where* was it?”, and “*In which context* was it?” (Saive *et al.*, 2013, 2014b, 2015). Our present study further developed this experimental protocol to investigate the effect of the sensory modality of the retrieval cue on EM retrieval. In our ecological laboratory protocol, virtual reality is used to favor EM encoding and retrieval (Smith, 2019). Participants discover events during a first-person exploration of rooms of a house, where boxes disposed at unique places simultaneously deliver an odor, a music and a face. Aiming to focus on sensory modality differences and to restrict potential influences of emotional aspects and semantic processes, odors, music and faces were selected to be emotionally as neutral as possible and as less identifiable as possible. As for most real-life events, the encoding was not explicit, but participants freely explored the three rooms on three consecutive days, one room per day. On the following day, EM retrieval was triggered with either odors, music or faces, and participants performed a series of tests. EM was investigated with the evaluation of the probability to (1) recognize the memory cue among distractors, and in case of recognition, to remember (2) the room and then (3) the box in which the cue was initially encoded, therefore allowing for a detailed examination of the richness of episodic retrieval. Response times and physiological responses (breathing) were also measured during these tasks. At the end of the testing session, participants evaluated the features of the memory cues (pleasantness, emotional intensity, familiarity and complexity). These additional evaluations aimed to further investigate the influence of potential sensory modality differences on memory performance. Considering the previously studied characteristics for EM of odor, music and faces, our hypothesis was that, while odor recognition might be not as good as music recognition or face recognition, once the cue was recognized, the retrieval of its encoding context would be more accurate with odors than the other sensory cues. Our study has thus three main goals: 1) to validate and further elaborate our laboratory-ecological approach; 2) to investigate the effect of the sensory modality of the cue on EM retrieval; and 3) to investigate respiratory behavior associated with episodic retrieval.

2 Materials and Methods¹⁰

2.1 Ethics Statement

All participants provided written informed consent prior to participation. The protocol was approved by the national Institutional Review Board, according to French regulations for biomedical experiments with healthy volunteers [Ethical Committee of CPP IDF8 (April 25,2017), ID RCB: 2016-A01931-50].

2.2 Participants

Fifty-four healthy participants [31 women and 23 men; aged 22.8 ± 2.3 (mean \pm standard deviation)] consented to participate to this experiment and received 40 Euros in compensation. All participants reported normal senses of smell and no visual or auditory impairments. They were recruited through campus electronic mails and posters. Given the implicit nature of the experiment, they were told that the study aimed to investigate the perception of various environments involving pictures, music and odors. Three participants had been excluded because of technical issues. Twenty-eight participants reported to have on average $5.7 (\pm 3.2)$ years of formal musical training on an instrument.

2.3 Stimuli and materials

2.3.1 Multidimensional episodes

A software called EpisOdor was developed using Unity 5.2.2 (Unity Technologies, USA) by a local technical platform *NeuroImmersion* (CRNL). EpisOdor allowed the presentation of a virtual 3D house in a first-person view (non-immersive virtual reality) that can be actively explored using a trackball (Kensington, Redwood Shores, CA, USA). This house was composed of three rooms, a bedroom, an office and a living room (defining the *Which context* component), connected to a corridor by closable doors (Figure 1A). Each of these rooms was defined by its furniture (*e.g.*, chair, bed, table, piano) and decorative elements (*e.g.*, painting, carpet). In addition, three clickable boxes were placed in each room at specific locations (defining the *Where* component), for a total of nine boxes. They were highlighted by an arrow that appeared when participants were nearby. These boxes were relatively distantly located from each other within a given room and were arranged differently between each room. EpisOdor had two modes. In the encoding mode, participants started in the corridor and a unique door was opened. When participants entered the room and clicked on a box, the box

¹⁰ This study used two versions of a same experimental implementation (Version1, V1 and Version2, V2). Given that the two versions differed only slightly for the retrieval test (see 2.4.2), and that the main results did not differ between versions, the two versions have been combined in the presentation, unless otherwise noted.

opened and three unique stimuli were presented simultaneously (defining the *What* component): an odor, a music and a face (Figure 1B). In the retrieval mode, participants started in the corridor where they faced a closed box and the three doors were closed. After clicking on the corridor's box, the three doors opened and the participants can enter any room and click on any box in the rooms, but no stimuli were presented. The virtual 3D-house and the location of the boxes in each room was constant across participants, but the three stimuli assigned to each box in the encoding mode was randomly defined across participants.

2.3.2 Stimuli and apparatus

Eighteen stimuli of each sensory modality (Odor, Music, Face) were selected from behavioral pre-tests from a set of 24 items [16 participants, using subjective scales from 0 (minimum) to 10 (maximum)] based on their neutral valence (from extremely unpleasant to extremely pleasant; mean \pm standard deviation; Odor, 4.9 ± 1.5 ; Music, 6.34 ± 0.6 ; Face, 5.3 ± 0.9), and relatively low and comparable emotional intensity (from emotion being evaluated as not intense at all to extremely intense; Odor, 5.5 ± 0.8 ; Music, 5.0 ± 1.0 ; Face, 4.2 ± 0.7), familiarity (from not familiar at all to extremely familiar; Odor, 5.8 ± 1.5 ; Music, 4.3 ± 1.2 ; Face, 3.5 ± 0.6) and complexity (from not complex at all to extremely complex; Odor, 5.8 ± 0.7 ; Music, 4.7 ± 1.2 ; Face, 4.9 ± 0.7), in order to limit the potential effects of subjective sensory evaluations on episodic memory (Brédart and Barsics, 2012; Cornell Kärnekull *et al.*, 2015; Holland and Kensinger, 2010; Saive *et al.*, 2014b; Stalinski and Schellenberg, 2013; Tolin and DeLeagnas, 1973) and to favor the comparability of the stimuli between sensory modalities. The 18 stimuli of each sensory modality were divided into two subsets of nine stimuli (set 1 and set 2) matched for pleasantness, emotional intensity, familiarity and complexity (two-sample t-test, t 's ≤ 0.48 , p 's ≥ 0.63). Sets 1 and 2 were pseudo-randomly assigned as target or distractor items for each participant, so that each set was equally represented as target or distractor among them.

2.3.2.1 Odorants

Odorants consisting of essential oils, single or mixtures of monomolecular chemical compounds, and fragrances were used. Set 1 was composed of Carrot (Givaudan-Roure, Vernier, Suisse), Cis-3-hexenyl salicylate (Créations aromatiques, Neuilly-Sur-Seine, France), Daffodil (Givaudan-Roure), Lovely Ion (EmoSens, Lyon, France), Methyl octine carbonate (Créations aromatiques), Musk (Givaudan-Roure), Secret de Hammam (EmoSens), Stemone (Créations aromatiques) and Teck Lounge (EmoSens). Set 2 was composed of 9-decen-1-ol (Sigma-Aldrich®, Saint-Louis, MS, USA), Basil (Créations aromatiques), Bien-Être (EmoSens), Birch oil (Sigma-Aldrich), Citronellol (Sigma-Aldrich), Linalyl acetate (Sigma-Aldrich), Osmose (EmoSens), Tobacco (Givaudan-Roure) and Tomato (Givaudan-Roure). The undiluted

odorants were placed in a 10 ml U-shaped Pyrex[®] tubes (VS Technologies, Saint-Priest, France) filled with microporous substances.

The odorants were presented with a 20-channel computer-controlled olfactometer adapted from (Sezille *et al.*, 2013). This odor diffusion system was developed to synchronize odorous stimuli with breathing. The participants' nasal respiratory signals were acquired using a nasal canula and were used to trigger the odor stimulation through an airflow sensor. During odor stimulation, the olfactometer waited for the participants' subsequent expiration, allowing the odor to be perceived at the beginning of the subsequent inspiration. When this expiration was detected, an unodorized airflow was sent to one of the U-shaped odorous tubes. Odorized airflow and air carrier were sent to and mixed in a homemade mixing head made of polytetrafluoroethylene. It was connected to the nostrils through two Teflon tubes, fixed to the nasal canula, opening out under the nostrils. The olfactometer airflow rate was set between 2.6 and 5.0 l/min depending on the odorant physical intensity. The odorants were delivered over 6 s. The olfactometer was controlled by an in-house LabView software (National Instruments, Austin, TX, USA) and interacted with EpisOdor software to synchronize odor stimulation with box opening via a TCI-IP connection.

2.3.2.2 Musical Pieces

Musical clips were selected from the materials used in Vieillard *et al.* (2008) (Copyright, Bernard Bouchard, 1998) (http://www.peretzlab.ca/knowledge_transfer/). They were modified in MIDI (Digital Performer[®], MOTU, Cambridge, USA) aiming to reduce potentially evoked emotional features. The modifications included changes of tempo, mode (major/minor), and/or a few notes, depending on the clip. The musical clips were played with an acoustic piano timbre (Cubase[®], Steinberg Media Technologies, Hamburg Germany) and presented with EpisOdor at a comfortable loudness level with headphones. Their average duration was 7.70 ± 1.30 s. The two music sets were matched for their number of events (28.50 ± 7.83), tempo (number of pulses per min, 76.72 ± 19.35) and duration (7.44 ± 1.58 s) (two-sample t-test, t 's ≤ 1.25 , p 's ≥ 0.22).

2.3.2.3 Faces

Faces were selected from the Chicago Face Database (Ma *et al.*, 2015) (<https://chicagofaces.org/default/>). They were composed of nine female and nine male faces from various origins (Asian, Caucasian, African American, Latin American; Figure 1C). All faces were presented on a white background, their relative position and dimensions were similar, and all persons wore a grey sweatshirt. To limit their distinctiveness, they were turned into black and white and the most prominent details of each face were erased (spots, freckle; Photoshop[®], Adobe, Dublin, Ireland). The 2863 x 1718 pixels pictures were presented in jpeg

format inside the box cover, on the center of the screen, with EpisOdor. The duration of their presentation was adjusted to the music duration when presented simultaneously, and was of 6 s otherwise.

2.4 Experimental procedure

The experiment was divided into four sessions, one session per day on consecutive days (Figure 1D). The first three days were devoted to the encoding phase, and the last day was divided into three sessions: retrieval test, questionnaire, and stimuli evaluation. A night of sleep followed each of the encoding sessions to promote consolidation and to reduce interference (Abichou *et al.*, 2019; Maquet, 2001; Stickgold, 2005). Participants completed the four sessions at the same time of each day to limit the differential influences of internal states (hunger, satiety) on olfactory and cognitive processes between sessions (Jiang *et al.*, 2008; Plailly *et al.*, 2011). The participants were requested to breathe through their nose as normally as possible without consciously modifying their respiration.

2.4.1 Encoding

Each encoding session consisted of at least 10 min during which participants freely explored the room with the door open, one different room on each of the three days. The order of the rooms was counterbalanced across participants. Participants were encouraged to pay attention to each detail of the room, and were told to click at least five times on each box to be sure they sufficiently explored each stimulus. The fifth click was indicated to the participant with a change of the box's arrow color, from green to red. No instructions for memorization were given, but participants were informed that they would be questioned about their perception of the episodes on the fourth day. This procedure aimed to ensure free encoding similar to real-life situations.

2.4.2 Retrieval

The retrieval test took place in the same 3D virtual house, and was composed of 54 trials. Each trial started in the corridor, with the 3 rooms closed, in the center of which a closed box was placed (Figure 1E). Participants were told to click on the box. The click triggered the opening of the box and the presentation of a unimodal cue (Odor, Music or Face). Half of the memory cues had never been encountered during the encoding phase (27 distractors), and the others had been explored during the three preceding days (27 targets). The cues were presented pseudo-randomly with the constraint to avoid the presentation of two odorants, or three music or faces in a row. This avoided olfactory saturation, and any habituation to the stimulus type. Participants were first asked whether they recognized the cue as having been presented at the encoding phase or not (Recognition response; “*Do you recognize this*

stimulus? (Y/N)). The answer was given with right or left click on the trackball. In the case of a “No” response, the trial ended. Therefore, the cue diffusion stopped, participants were taken back into the corridor, and another closed box containing another cue was placed in front of them. In the case of a “Yes” response, in V1, the cue diffusion stopped, the box closed, and the corridor appeared with all doors opened. Participants were asked to move into a room in order to click on the box in which they previously encountered the cue. A click on a box ended the trial (Episodic response). In V2, the stimulus diffusion was maintained until participants had given their response to the question about whether they remembered the context (the room and the box) where they previously encountered the cue (Subjective recollection response; *“Do you remember the context? (Y/N)”*), allowing for metacognitive knowledge evaluation. Regardless of their response, the doors opened and participants had to move into a room to click on the box in which they previously encountered the cue, and a click on a box ended the trial (as in V1).

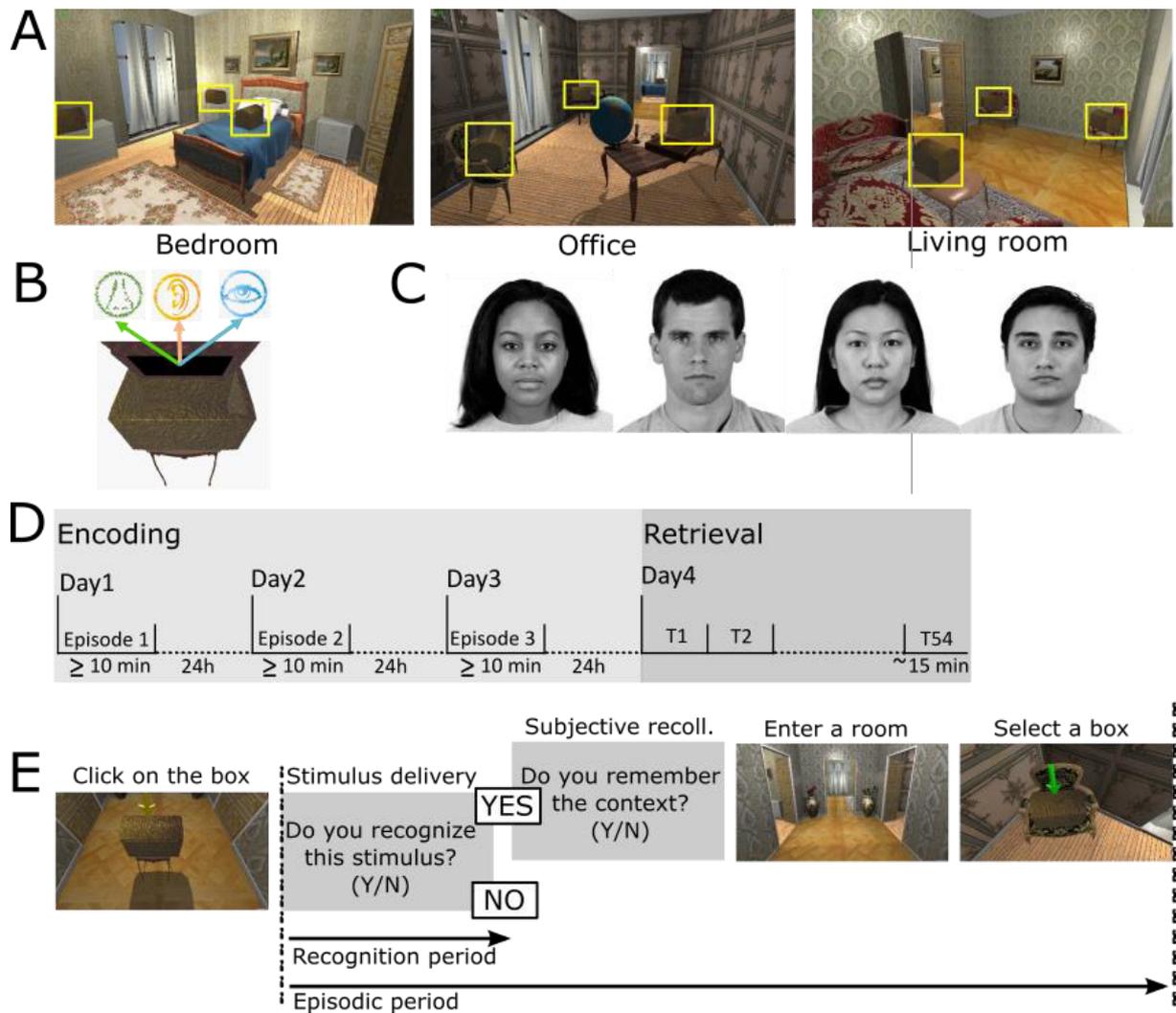


Figure 1. Experimental protocol. (A) The three rooms of the virtual house. Yellow rectangles represent the box in which the stimuli were presented. (B) Detail of a box, containing an odor, a music, and a face. (C) Example of face stimuli. (D) The temporal course of the encoding and the retrieval sessions. During the encoding, participants discovered one episode a day over three days. On the fourth day, the memory of the episodes was tested. (E) Example of a trial. Participants click on the box and a stimulus, either an odor, a music or a face was presented. The participants must first recognize whether the cue was a target or a distractor and, in case of a ‘Yes’ response, for V2 only, inform us about its subjective recollection and then enter a room to select a box. In V1, in case of a ‘Yes’ response, participants must directly enter a room and select a box.

2.4.3 Sensory evaluations

At the end of the experiment, participants were required to rate the 54 stimuli (targets and distractors of all modalities) in terms of pleasantness (unpleasant – neutral – pleasant), emotional intensity (very weak – very intense), familiarity (unknown – very familiar), and complexity (very simple – very complex) using non-graduated scales presented with an in-house LabView software. The pleasantness scale was divided into two equal parts by a

“neutral” value separating the ratings of unpleasantness and pleasantness. Each unimodal stimulus was presented for about 6 s and participants had been given 5 s to respond to each question. Participants were trained with a short practice session consisting of one item per sensory modality.

2.5 Data analyses

2.5.1 Episodic memory performance

Episodic memory retrieval has been divided into several steps. It first necessitates to recognize accurately from all cues presented (step “Cue”) the target cue (step “Cue+”) and then to enter the accurate room (step “Room+”) and to choose the accurate box (step “Box+”), in which the cue had been presented during encoding (Figure 2A). Note that conditions were interdependent, such that Room+ implied Cue+, and Box+ implied Room+. Based on the ability of the participants to reach each step or not, memory episodic responses were organized into categories. For recognition responses, four response categories were defined from the experimental conditions (target vs. distractor) and the participants' behavioral responses (“Yes” vs. “No”) based on the signal detection theory (Lockhart and Murdock, 1970): *Hit* and *Miss* occurred when the target items were correctly recognized or incorrectly rejected, respectively, and correct rejection (*CR*) and false alarm (*FA*) occurred when the distractor items were correctly rejected or incorrectly recognized, respectively. For episodic responses, based on the room entered and the box selected, three response categories were defined: *Room-Box-* when the wrong room was entered (and thus a wrong box was selected), corresponding to a completely incorrect episodic response, *Room+Box-* when the accurate room was entered, but an inaccurate box was selected, corresponding to a partially correct episodic response, and *Room+Box+* when the participants entered the accurate room and selected the accurate box, corresponding to a complete correct episodic response.

EM performance was computed as the probability to reach each episodic step: (1) recognize the memory cue among distractors (*from Cue to Cue+*; 1 response out of 2, “Yes” or “No”; *i.e.*, 50 % chance to pick the correct answer), and in case of recognition, to remember (2) the room (*from Cue+ to Room+*; 1 room out of 3; *i.e.*, 33 % chance to enter the accurate room) and then (3) the box (*from Room+ to Box+*; 1 box out of 3; *i.e.*, 33 % chance to click on the accurate box) in which the cue was initially encoded. EM global performance was evaluated in computing the probability to fully remember the episodes (4) from the cue presentation (*from Cue to Box+*; $1/2 * 1/3 * 1/3$; *i.e.*, 5 % chance), and (5) from the cue recognition (*from Cue+ to Box+*; $1/3 * 1/3$, *i.e.*, 11 % chance) (Figure 2A). All these probabilities were 1) analyzed as a function of Sensory Modality (odor, music, face) and Version (V1, V2) and 2) compared to chance levels.

Additionally, recognition performance was resumed with two metrics: The memory score (d') reflects the participant's ability to discriminate between target and distractor items (the higher the d' , the better the performance), and response bias score (c) reflects participant's tendency to give more "Yes" (positive score) or "No" (negative score) response regardless of experimental condition.

2.5.2 Subjective recollection

Subjective recollection responses in V2 were analyzed through the probability to give a "Yes" response (1 "Yes" or "No" response out of 2, *i.e.*, 50 % chance). It was analyzed as a function of Sensory Modality (odor, music, face) and of subsequent episodic memory responses (Room-Box-, Room+Box-, Room+Box+).

2.5.3 Response times

Response times were computed for cue recognition, subjective recollection (in V2 only) and episodic retrieval response categories. The response times corresponded to the durations between the presentation of the cue (or the first inspiration after the cue in case of odors) and 1) the "Yes/No" recognition response, 2) the "Yes/No" response for the subjective recollection, and 3) the click on the box for the episodic retrieval. Outlier response times were removed by setting aside data that were above the 9th decile of the mean computed by Participant and Sensory Modality. Response times were analyzed 1) as a function of Memory response category (recognition period: Hit, Miss, CR, FA; episodic period: Room-Box-, Room+Box-, Room+Box+), Sensory Modality (odor, music, face) and Version (V1, V2) and for recognition, 2) as a function of recollection response and Sensory modality for recollection period, 3) as a function of subsequent episodic memory responses (Room-Box-, Room+Box-, Room+Box+).

2.5.4 Breathing parameters

Breathing parameters consisted of respiratory frequency, and inspiration amplitude and volume. They were computed for the same conditions and within the same boundaries than the one defined for RTs analyses. Breathing parameters were analyzed as a function of Memory Response categories (recognition period: Hit, CR; episodic period: Room-Box-, Room+Box-, Room+Box+), Sensory Modality (odor, music, face) and Version (V1, V2).

2.5.5 Encoding variables

For each cue, the Number of Presentation (*i.e.*, number of clicks on the box) and their day of presentation (Encoding day; day1, day2, day3) were computed. The relationship between the encoding variables and the episodic memory performance was investigated by analyzing the effect of Number of Presentation and Encoding day (day1, day2, day3) on each step of interest

of the episodic memory process (from Cue to Cue+, from Cue+ to Room+, from Room+ to Box+, from Cue to Box+, from Cue+ to Box+).

2.5.6 Sensory evaluations

Pleasantness, emotional intensity, familiarity, and complexity ratings of the stimuli were *a posteriori* transformed into scores from 0 to 10. Due to technical issues, some data were missing. They were analyzed as a function of Sensory Modality (odor, music, face). The relationship between the sensory evaluations of the cues and the episodic memory performance was investigated by analyzing the effect for each sensory evaluation (Pleasantness, Emotional intensity, Familiarity, Complexity) on each step of interest of the episodic memory process (from Cue to Cue+, from Cue+ to Room+, from Room+ to Box+, from Cue to Box+, from Cue+ to Box+).

2.6 Statistical analyses

2.6.1 Linear and generalized linear mixed effect models

2.6.1.1 General method

Except for breathing data (2.6.2), the analyses were conducted with linear and generalized linear mixed effect models (LMM and GLMM, respectively), which contain both fixed and random effects. Fixed effects are the effects of interest, whereas random effects represent levels that randomly varies across population that are expected to be controlled for. Here, Sensory Modality was a fixed effect as the major goal of this study was to determine how memory performance varied within its three levels (Odor, Music, Face). Participant and/or Item (sensory cues) were accounted for random effect to explain between-Participants or between-Item variability without investigating them [for more explanations about fixed and random effects, see (Brauer and Curtin, 2018; Brown, 2021; Singmann and Kellen, 2019)]. These models can represent data following Gaussian variation (LMM) as well as other kind of variations, from gamma to binomial (GLM). In addition to the fact that these methods are more frequently now used in biology and psychology research (Brown, 2021; DeBruine and Barr, 2021; Harrison *et al.*, 2018; Meteyard and Davies, 2020), this choice had been motivated by three main advantages. First, they allowed for analyzing the effect of Sensory Modality on memory accuracy by taking into account its binomial nature (Jaeger, 2008; Meteyard and Davies, 2020). Second, they allowed modeling more than one random effects (non-controlled environmental variability), and thus to render our results more powerful and generalizable (Jaeger, 2008; Brown, 2021; DeBruine and Barr, 2021). Third, they allowed integrating continuous variables as fixed effect. For example, it was possible here to investigate the effect

of stimulus evaluation on accuracy without discretizing the evaluation variable. Mixed models are also more robust to missing data and unbalanced design (Brown, 2021).

In the current study, mixed models were fitted with the “mixed” function from the *afex* library (Bates *et al.*, 2015; Singmann *et al.*, 2021). These mixed models were fitted by the Laplace approximation method. The global procedure used to construct models was the following: Unless specified, the maximal model justified by the design and the number of datapoints was constructed first (Barr *et al.*, 2013). In case of singular fit or convergence errors, random effects were selected (reduced) by comparing nested models with likelihood ratio test (LRT) and a backward selection heuristic, so the model complexity was reduced until the statistical test was significant (same method as in Matuschek *et al.*, 2017) or no more errors were noted. As the experiment had been constructed with a within-participant design, the random Participant intercept was never dropped. LMM were checked using package *performance* (Lüdtke *et al.*, 2021) from the *easystat* collection (Makowski *et al.*, 2020). GLMM were checked using the *DHARMA* package (Hartig, 2021) and *p*-values were computed by LRT. *P*-values were computed using the package *lmerTest* (Kuznetsova *et al.*, 2015) by the Kenward-Roger approximation with type 2 tests when the model did not contain interactions, and type 3 tests when it did.

2.6.1.2 Effect of sensory modality of the cue on episodic memory performance

The first step of episodic memory, which corresponds here to recognition responses, has been analyzed by adapting the signal detection theory framework to GLMM analysis (Fawcett and Ozubko, 2016; Wright and London, 2009; Wright *et al.*, 2009), using sum contrast coding so that the intercept represents reference modalities. A probit mixed model (estimated using maximum likelihood or ML, and the base optimizer) was fitted to predict the probability to answering “Yes” with a 3 Sensory Modality (odor, music, face) x 2 Item Type (target, distractor) x 2 Version (V1, V2) mixed design. Item Type and Version were centered, applying the following dummy coding: distractors were coded as -0.5 and targets were coded as 0.5, V1 was coded as -0.5 and V2 was coded as 0.5. For Sensory Modality, the reference level was “face”, and for Item Type and Version, it was 0. This coding allowed us to compute d' and c scores for each sensory modality using the estimates calculated by the model. The Item Type estimate represented the propensity to discriminate between old and new face, as face was the reference level. It was thus a measure of face's d' . For odor and music, this value was equal to the value for face plus the estimate of the interaction between Sensory Modality and Item Type for odor and music (*i.e.*, for odor, “ d'_{face} ” estimate + “Sensory Modality [odor] * Item Type” estimate). As the intercept represents the propensity to say “Yes” for a face with Item Type and Version at their mean level, its probit-transformed estimates were equal to the bias

score (*c*) for faces. Bias scores for odor and music were then calculated by adding their respective coefficient to the intercept value (*i.e.*, for odor, “*C_{face}*” estimate + “Sensory Modality [odor] * Item Type” estimate). This model allowed us also to compute the probability to reach Cue+ from Cue. The maximal random effect structure for this model was: (Item Type * Sensory modality | Participant) + (Item Type | Item).

For episodic step of interest following accurate cue recognition (from Cue+ to Room+, from Room+ to Box+, from Cue to Box+, from Cue+ to Box+), a logistic mixed model was modeled in order to predict the probability to succeed in each level with a 3 Sensory Modality (odor, music, face,) x 2 Version (V1, V2) mixed design using logit link and estimated with ML and base optimizer. The maximal random effect structure was the same for all models: (Sensory Modality | Participant) + (Version | Item).

2.6.1.3 Subjective recollection

The maximal logit mixed model was constructed with Sensory Modality (odor, music, face) and Episodic Score (Room-, Box-, Box+) and their interaction as fixed effect. The maximal random effect structure was: (Sensory Modality * Episodic_Score | Participant) + (Episodic_Score | Item).

2.6.1.4 Response times

Response times were log transformed after an inconclusive try to fit gamma or inverse gaussian models to take into account the right skewness response time distribution (De Boeck and Jeon, 2019; Lo and Andrews, 2015). All models were fitted with response category (for recognition period: Hit, Miss, CR, FA; for episodic period: Room-Box-, Room+Box-, Room+Box+), Sensory Modality (odor, music, face), Version (V1, V2), and all double interactions as fixed effects. The maximal random effect structure was: (Sensory Modality | Participant) + (1 | item). It was then reduced to minimize convergence errors. Estimations were computed with the basal (*nloptwrap*) optimizer and REML, unless specified.

2.6.1.5 Sensory evaluations

For each sensory evaluation (pleasantness, emotional intensity, familiarity, and complexity) a linear mixed model was fitted (estimated using restricted maximum likelihood, or REML, and *nloptwrap* optimizer) with a 3 Sensory Modality (odor, music, face) x 2 Version (V1, V2) mixed design. The maximal random effect structure was: (Sensory Modality | Participant) + (1 | Item) (see Singmann and Kellen, 2019 for more details). Random effect structure for all the final models and transformations of dependent variable differed for each sensory evaluation.

2.6.1.6 Effect of encoding variables on memory performance

For each episodic step of interest (from Cue to Cue+, from Cue+ to Room+, from Room+ to Box+, from Cue to Box+, from Cue+ to Box+), a model was fitted to predict the probability to succeed in each level, with as fixed effects: Number of Presentation, Encoding Day, Version, the four second-order interactions between Sensory Modality or Version with Number of Presentation or Encoding Day, and the two third-order interactions between Sensory Modality, Version and Number of Presentation or Encoding Day. The maximal model of fixed effect with 3 Encoding day (day1, day2, day3) x 3 Sensory Modality (odor, music, face) x 2 Version (V1, V2) x 1 Number of Presentation was not fitted to have a powerful model while being able to test the effect of encoding variables. For recognition response, maximal random effect structure was: (Number of Presentation * Encoding Day * Sensory Modality | Participant) + (Number of Presentation + Encoding Day | Item). For other episodic responses, random interactions and Sensory Modality were not fitted in the maximal random model, as they were too few observations for such a complex model. The maximal random effect structure was: (Number of Presentation + Encoding Day | Participant) + (Number of Presentation + Encoding Day | Item).

2.6.1.7 Effect of sensory evaluations on memory performance

Exploratory analyses were conducted for each episodic step of interest (from Cue to Cue+, from Cue+ to Room+, from Room+ to Box+, from Cue to Box+, from Cue+ to Box+), to test the effect of Sensory Evaluations (Pleasantness, Emotional Intensity, Familiarity, Complexity) on the probability to reach that step. This type of analyses was chosen because we did not have *a priori* hypothesis on results, as stimuli were selected to be as similar as possible. We first constructed the maximal fixed and random effect models. These models included Sensory Modality, quadratic Pleasantness, Emotional Intensity, Familiarity, Complexity, and all second order interactions between Sensory Evaluation and Sensory Modality. To avoid over-fitting, random effect structure included only by-Participant and by-Item intercepts, by-Item intercept being removed in case of fitting errors. Fixed effects were then selected to avoid model overfitting. The optimal fixed effects structure was determined by an automated AICc (Akaike Information Criterion corrected for small sample size) backward selection technique from the *MuMIN* library version 1.43.17 (Bartoń, 2020). This technique allows computing AICc score to each sub-model and to select the best model, with the lowest AICc.

2.6.2 Breathing

During both recognition and episodic periods, breathing parameters (breathing frequency, inspiration amplitude and inspiration volume) were analyzed with two-way analysis of variances (ANOVAs) for repeated measures using the *aov_car* function from the *afex* library with Sensory Modality (odor, music, face) and Response Category (recognition: Hit, CR; episodic: Room-Box-, Room+Box-, Room+Box+) as within-participants factors, and Version (V1, V2) as between-participants factor. For recognition period, due to lack of inaccurate rejection (Miss) and recognition (FA) response categories, only Hit and CR response categories were considered. For both recognition and episodic periods, due to data unbalanced between participants, the effects of episodic responses on breathing data were tested separately for the three sensory modalities. When Mauchly test for Sphericity was significant, Greenhouse-Geisser correction was applied. When ANOVA assumptions were violated, Friedman tests were computed and followed by pairwise tests of Wilcoxon with Bonferroni correction.

2.6.3 General statistical information

Statistical analyses were performed in R studio with R 4.0.4 (R Development Core Team, 2021), with orthogonal sum-to-zero contrasts set in the default settings. Mixed models and ANOVAs were fitted with the *afex* library version 0.28-1 and based on *lme4* package. When fixed effects were significant, *post-hoc* Bonferroni comparisons were run to determine significant differences among the estimated marginal means (EMM) with *emmeans* package version 1.5.5-1 (Lenth *et al.*, 2020). The same package was used to test the difference with chance value. The significance level was set to $p < .05$, and a p -value between 0.05 and 0.09 was referred to as a tendency. For clarity purpose, only tests addressing the main goal of the present paper were reported. Similarly, when interaction between effects was statistically significant, effects of the main factors were not reported. When interaction between variables were tested, the trend of the slope was reported. Data were described with estimated marginal means and its standard error of the mean (EMM \pm SEM).

3 Results

3.1 Memory performance

Memory performance are displayed in Figure 2B.

3.1.1 From Cue to Cue+

Hit (*i.e.*, accurate recognition of target cues) and FA (*i.e.*, inaccurate recognition of distractor cues) recognition responses were analyzed with the [(Sensory Modality + Item Type II

Participant) + (Item Type | Item)] random effect structure (conditional $R^2 = 0.72$, marginal $R^2 = 0.72$). A significant interaction between Sensory Modality and Item Type ($\chi^2(2) = 57.66$, $p < .001$) showed that the probability for a Hit was significantly higher for music (EMM \pm SEM, 98.5 ± 0.8 %) and face (95.0 ± 1.5 %) than for odor (79.90 ± 3.8 %; z 's ≥ 4.45 , p 's $< .001$), and tended to be higher for music than faces ($z = 2.22$, $p = .08$), and that the probability for an FA response was higher for odor (23.11 ± 4.1 %) than for both music (4.51 ± 1.6 %) and face (2.36 ± 0.9 %; z 's ≥ 4.71 , p 's $< .001$). In other words, odor cues evoked less accurate recognition and more inaccurate recognition than music and face cues. Moreover, the probability to obtain Hit and FA recognition responses differed significantly from random responses (50 %) suggesting good recognition performance of cues from all three sensory modalities (z 's ≥ 6.53 , p 's $< .001$). Memory score d' was estimated at 1.57 ± 0.26 for odors, 3.86 ± 0.34 for music, and 3.63 ± 0.23 for faces. Bias score c was estimated at 0.05 ± 0.14 for odors, 0.25 ± 0.15 for music, and -0.17 ± 0.10 for faces.

3.1.2 From Cue+ to Room+

The model computing the probability to retrieve the accurate room after accurately recognizing a target cue was fitted with maximal random-effect structure (conditional $R^2 = 0.15$, marginal $R^2 = 0.11$). This probability was significantly influenced by the Sensory Modality of the cue ($\chi^2(2) = 44.47$, $p < .001$), such that it was higher for face (71.6 ± 3.45 %) than for both music (34.8 ± 2.46 %) and odor (51.5 ± 2.99 %; z 's ≥ 4.49 , p 's $< .001$), and higher for odor than for music ($z = -4.31$, $p < .001$). Moreover, this probability was significantly higher than chance (33 %) for odor and face (z 's ≥ 6.09 , p 's $< .001$), but not for music ($z = 0.58$, $p = 1.00$).

3.1.3 From Room+ to Box+

The model computing the probability to retrieve the accurate box after entering the accurate room was fitted with [Sensory Modality | Participant] random-effect structure (conditional $R^2 = 0.11$, marginal $R^2 = 0.08$). This probability was significantly affected by the Sensory Modality of the cue ($\chi^2(2) = 16.35$, $p < .001$), such that it was higher for faces (73.0 ± 4.29 %) than for odors (52.2 ± 4.33 %) and music (44.4 ± 4.39 %; z 's ≥ 3.34 , p 's $< .001$). The probability for odors and for music did not significantly differ ($z = 1.19$, $p = .70$). Moreover, this probability was significantly higher than chance (33 %) for cues of all sensory modalities (z 's ≥ 2.52 , p 's $< .05$).

3.1.4 From Cue to Box+

The model computing the probability to accurately retrieve the complete episode from the presentation of a cue (distractor or target) was fitted with (Sensory Modality | Participant) random effect structure (conditional $R^2 = 0.16$, marginal $R^2 = 0.12$). This probability was

affected by the Sensory Modality of the cue ($\chi^2(2) = 34.17, p < .001$), such that it was higher for faces ($44.5 \pm 4.42\%$) than odors ($19.2 \pm 2.37\%$) and music ($14.5 \pm 1.90\%$; z 's $\geq 5.69, p$'s $< .001$). The probability was not significantly different for odors and music ($z = 1.46, p = .43$). Moreover, this probability was significantly higher than chance (5%) for cues of all sensory modalities (z 's $\geq 4.70, p$'s $< .001$).

3.1.5 From Cue+ to Box+

The model computing the probability to retrieve the complete episode after accurately recognizing a target cue was fitted with (Sensory Modality | Participant + Version | Item) random effect structure (conditional $R^2 = 0.17$, marginal $R^2 = 0.13$). This probability was significantly affected by the Sensory Modality of the cue [$\chi^2(2) = 34.96, p < .001$], such that it was higher for faces ($48.2 \pm 4.53\%$) than for odors ($26.0 \pm 2.91\%$) and for music ($14.8 \pm 1.97\%$), and higher for odors than music (z 's $\geq 2.98, p$'s $< .01$). Moreover, this probability was significantly higher than chance (11%) for odors and faces (z 's $\geq 5.11, p$'s $< .001$), but not music ($z = 1.89, p = .17$).

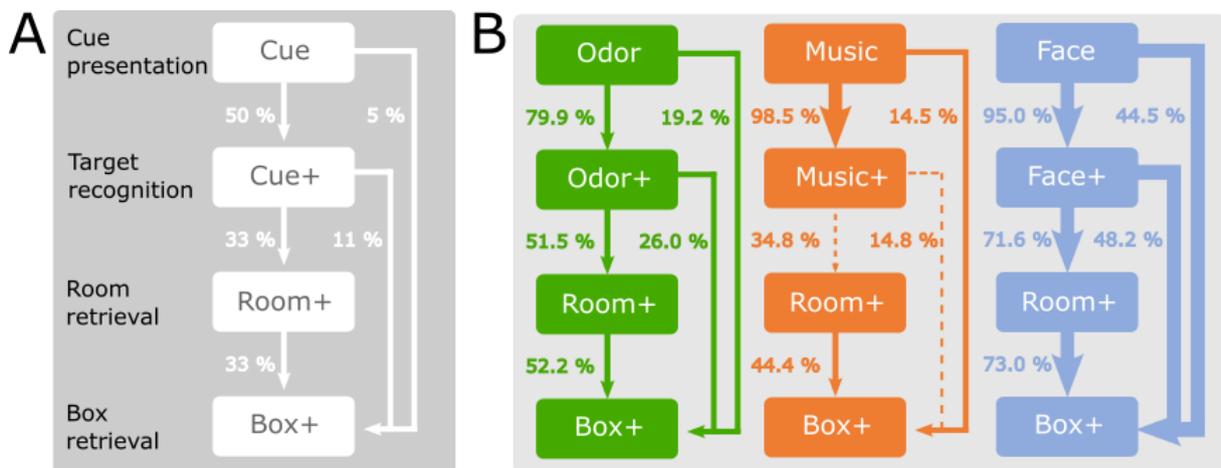


Figure 2. Episodic memory performance. A. Experimental conditions with chance-level value of probability to go from one step to another. B. Sensory-modality specific probability to go from one step to another. Differences of arrow thickness among a same step represent significant difference of probability between sensory modalities. Solid lines represent significant probability; dashed lines represent non-significant probability (see supplementary materials for non-modality specific memory performance).

3.2 Subjective recollection

In V2, metacognition was evaluated, as participants were asked whether they felt they could remember the entire episode. The probability to answer “Yes” was fitted with a model containing (Sensory Modality * Episodic Score | Participant) random effect structure

(conditional $R^2 = 0.46$, marginal $R^2 = 0.23$). The probability to answer “Yes” significantly differed with Sensory Modality (Figure 3A; $\chi^2(2) = 41.52$, $p < .001$). It was higher for faces ($82.38 \pm 5.57\%$) than for both odors ($48.06 \pm 9.14\%$) and music ($32.23 \pm 7.42\%$; z 's ≥ 3.85 , p 's $< .001$). Moreover, this probability was significantly above chance (50%) for faces ($z = 5.82$, $p < .001$), below chance for music ($z = 2.40$, $p < .05$), but not different than chance for odors ($z = 0.21$, $p = 1.00$). The probability to answer “Yes” was also significantly influenced by Episodic Score (Figure 3B; $\chi^2(2) = 7.22$, $p < .05$). It was higher before an accurate episodic response (Room+Box+; $70.85 \pm 8.40\%$) than before an inaccurate episodic response (Room-Box-; $44.45 \pm 8.13\%$; $z = 2.82$, $p < .05$), but not before a partially accurate episodic response (Room+Box-; $51.41 \pm 9.95\%$; $z = 1.42$, $p = .28$). Moreover, this probability was above chance level (50%) when followed by complete episodic response only (Room+Box+, $z = 2.48$, $p < .05$; Room-Box- & Room+Box-, z 's < 0.68 , p 's = 1.00).

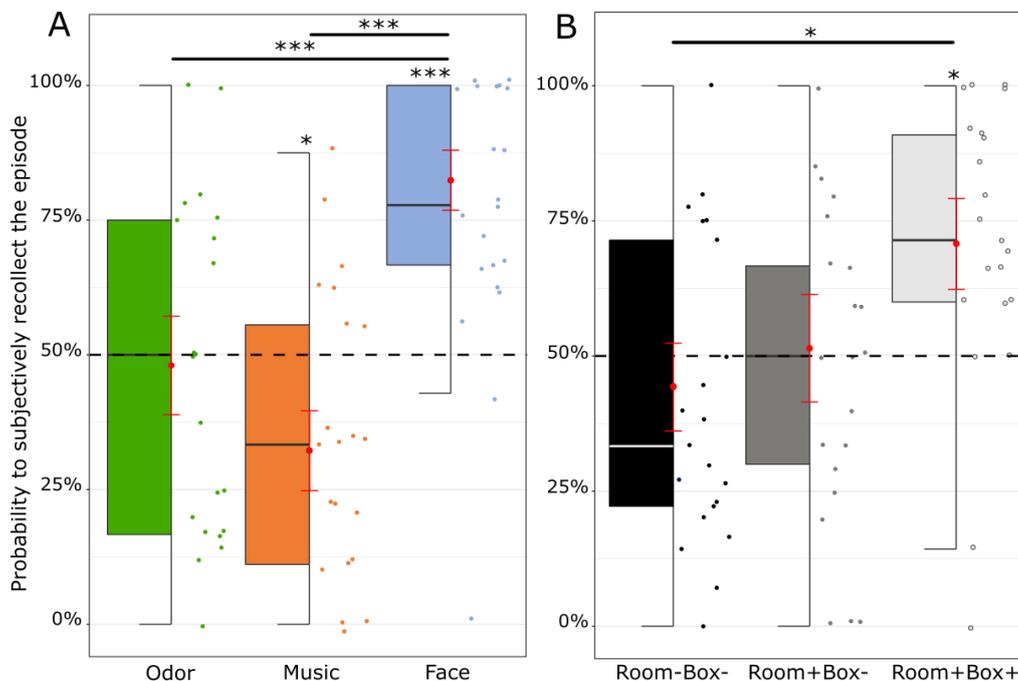


Figure 3. Subjective recollection responses. Probability to give a “Yes” response as a function of A. sensory-modality of the cue, B. subsequent episodic memory response. Individual raw data are represented with dots. The distribution of data is displayed with boxplot (minimum, first quartile, median, third quartile, maximum) in black. The model estimated means and their dispersion (SEM) are represented in red. The dashed horizontal line indicates the random level. * $p < .05$; *** $p < .001$.

3.3 Response times

3.3.1 Recognition period

The recognition response times were fitted with maximal random-effect structure model (conditional $R^2 = 0.54$, marginal $R^2 = 0.28$). Significant effect of the interaction between Sensory Modality and Response Categories ($F_{(6,2555)} = 12.18$, $p < .001$) revealed that correct responses were faster than incorrect ones for both target and distractor cues (Figure 4A): Hit responses were faster than Miss responses for odors (Hit, 4.46 ± 0.28 s; Miss, 6.20 ± 0.45 s), music (Hit, 4.39 ± 0.28 s; Miss, 6.34 ± 0.82 s), and faces (Hit, 2.57 ± 0.14 s; Miss, 4.17 ± 0.39 s; t 's(2553) ≥ 3.12 , p 's $< .05$), and CR responses were faster than FA responses for odors (CR, 4.78 ± 0.30 s; FA, 5.58 ± 0.39 s), and faces (CR, 2.24 ± 0.13 s; FA, 5.10 ± 0.58 s; t 's(2579) ≥ 3.17 , p 's $< .01$), but not for music (CR, 4.90 ± 0.31 s; FA, 5.90 ± 0.58 ; $t_{(2555)} = 2.27$, $p = .14$). Among correct responses, CR responses were faster than Hit responses for faces ($t_{(2498)} = 4.75$, $p < .001$), Hit responses were faster than CR responses for music ($t_{(2499)} = 3.67$, $p < .01$), and no differences were found for odors ($t_{(2528)} = 2.02$, $p = .26$). Among incorrect responses, RT did not differ significantly between Miss and FA for all three sensory modalities (t 's(2572) ≤ 1.75 , p 's $\geq .48$).

For Hit responses, the response times as a function of subsequent episodic memory responses were fitted with maximal random-effect structure model (conditional $R^2 = 0.51$, marginal $R^2 = 0.23$). There was a significant interaction between Sensory Modality and Version ($F_{(2,49)} = 6.32$, $p < .01$), revealing that music induced faster recognition responses in V2 (3.77 ± 0.36 s) than in V1 (4.88 ± 0.40 s; $t_{(50)} = 2.21$, $p < .05$). An effect of Subsequent Episodic Responses tended to be significant (Figure 4B; $F_{(2,1097)} = 3.00$, $p = 0.050$), suggesting that Hit responses were faster when followed by accurate episodic responses (Room+Box+, 3.55 ± 0.18 s) than inaccurate episodic responses (Room-Box-, 3.84 ± 0.18 s; $t_{(1103)} = 2.44$, $p < .05$), but that they were not significantly different than Hit responses followed by a partial episodic response (Room+Box-, 3.70 ± 0.19 s; $t_{(1097)} = 1.14$, $p = 0.77$). Similarly, inaccurate and partial episodic responses were not different between each other ($t_{(1090)} = 1.13$, $p = 0.81$).

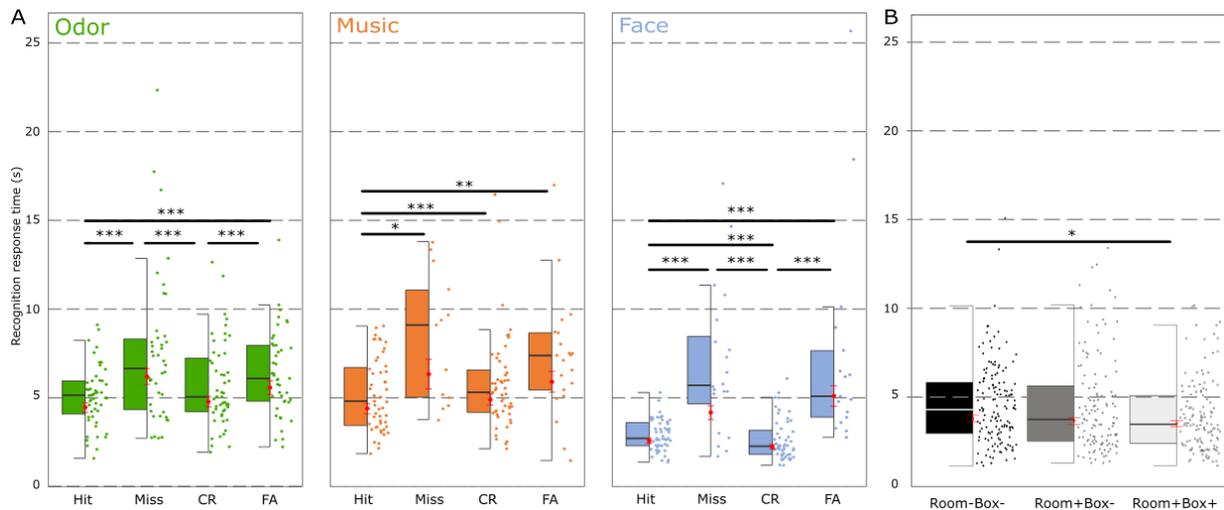


Figure 4. Recognition response times. Time (in s) to give a response (from cue presentation to recognition response) A. as a function of recognition response category for Odor, Music and Face, B. as a function of subsequent episodic response category for Hit response only. Individual raw data are represented with dots. The distribution of data is displayed with boxplot (minimum, first quartile, median, third quartile, maximum) in black. The model estimated means and their dispersion (SEM) are represented in red. * $p < .05$, ** $p < .01$, *** $p < .001$.

3.3.2 Subjective recollection

The effects of recollection response (Yes, No) and Sensory modality on the subjective recollection response times were first investigated with both distractors and target stimuli. The subjective recollection response time was fitted with a model using [(Sensory Modality + Recollection responses || Participant) + (1 | Item)] as random effects ($R^2 = 0.17$, marginal $R^2 = 0.16$). A significant interaction between Sensory Modality and Recollection Response ($F_{(2,280)} = 7.91$, $p < .001$) revealed that “Yes” responses for subjective recollection were faster for faces (6.23 ± 0.56 s) than for music (8.70 ± 0.86 s; $t_{(113)} = 4.64$, $p < .001$). No differences were observed with odor recollection response times (7.26 ± 0.73). Participants also responded “Yes” faster than “No” after face cues (“No”, 10.28 ± 1.13 s; $t_{(112)} = 5.48$, $p < .001$) and odor cues (No, 9.82 ± 0.95 s; $t_{(94)} = 3.41$, $p < .001$), but not a music (“No”, 9.32 ± 0.84 s; $t_{(79)} = 0.84$, $p = 0.36$).

The effects of the Sensory modality and the subsequent episodic score (Room-Box-, Room+Box-, Room+Box+) on the subjective recollection response times were then investigated in Hit and Yes responses only. The subjective recollection response time was fitted with a model using [(Sensory Modality || Participant) + (1 | Item)] as random effects ($R^2 = 0.50$, marginal $R^2 = 0.08$). A significant main effect of Sensory Modality (Figure 5; $F_{(2,65)} = 8.14$, $p < .001$) revealed that responses after a subjective recollection for faces (6.49 ± 0.65 s) were faster than for music (8.98 ± 1.02 s; $t_{(66)} = 3.95$, $p < .001$) and no difference with odor response time were observed (7.66 ± 0.84 s; t 's < 2.15, p 's > .10).

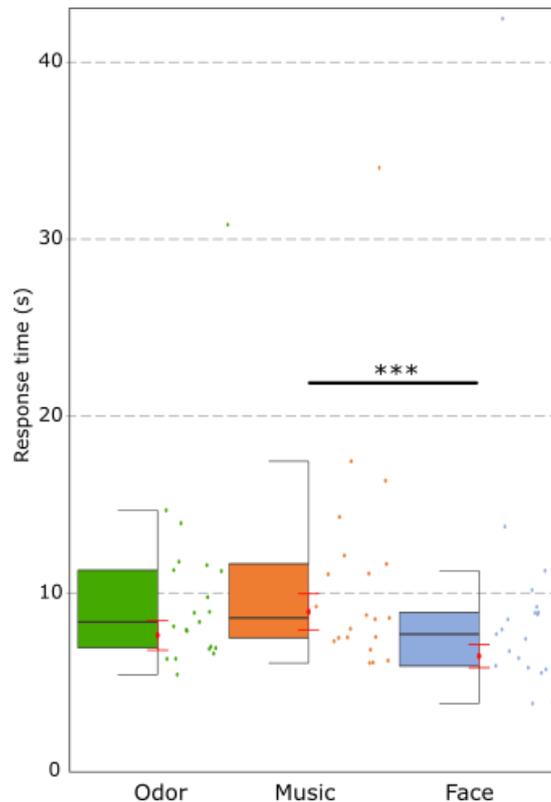


Figure 5. Subjective recollection response times. Time (in s) to give a response as a function of sensory-modality of the cue for subjectively recollected episodes. Individual raw data are represented with dots. The distribution of data is displayed with boxplot (minimum, first quartile, median, third quartile, maximum) in black. The model estimated means and their dispersion (SEM) are represented in red. The dashed horizontal line indicates the random level. *** $p < .001$.

3.3.3 Episodic period

The episodic response times were fitted with a model containing (1 | Participant) as random effect (conditional $R^2 = 0.31$, marginal $R^2 = 0.05$). A significant interaction between Sensory Modality and Episodic Response (Figure 6; $F_{(4,1145)} = 3.48$, $p < .01$) revealed that correct episodic responses were faster than incorrect episodic responses for odors (Room+Box+, 16.88 ± 0.95 s; Room-Box-, 20.78 ± 1.00 s) and faces (Room+Box+, 15.89 ± 0.74 s; Room-Box+, 19.71 ± 1.01 s; t 's ≥ 3.81 , p 's $< .001$), but not music (Room+Box+, 19.88 ± 1.23 s; Room-Box-, 19.94 ± 0.86 s; $t_{(1144)} = 0.05$, $p = 1.00$). Similarly, correct episodic responses were faster than partial episodic ones for odors (Room+Box-, 22.76 ± 1.31 s) and faces (Room+Box-, 18.53 ± 1.07 s; t 's ≥ 2.81 , p 's $< .05$), but not music (Room+Box-, 20.32 ± 1.17 s; $t_{(1142)} = 0.32$, $p = 1.00$).

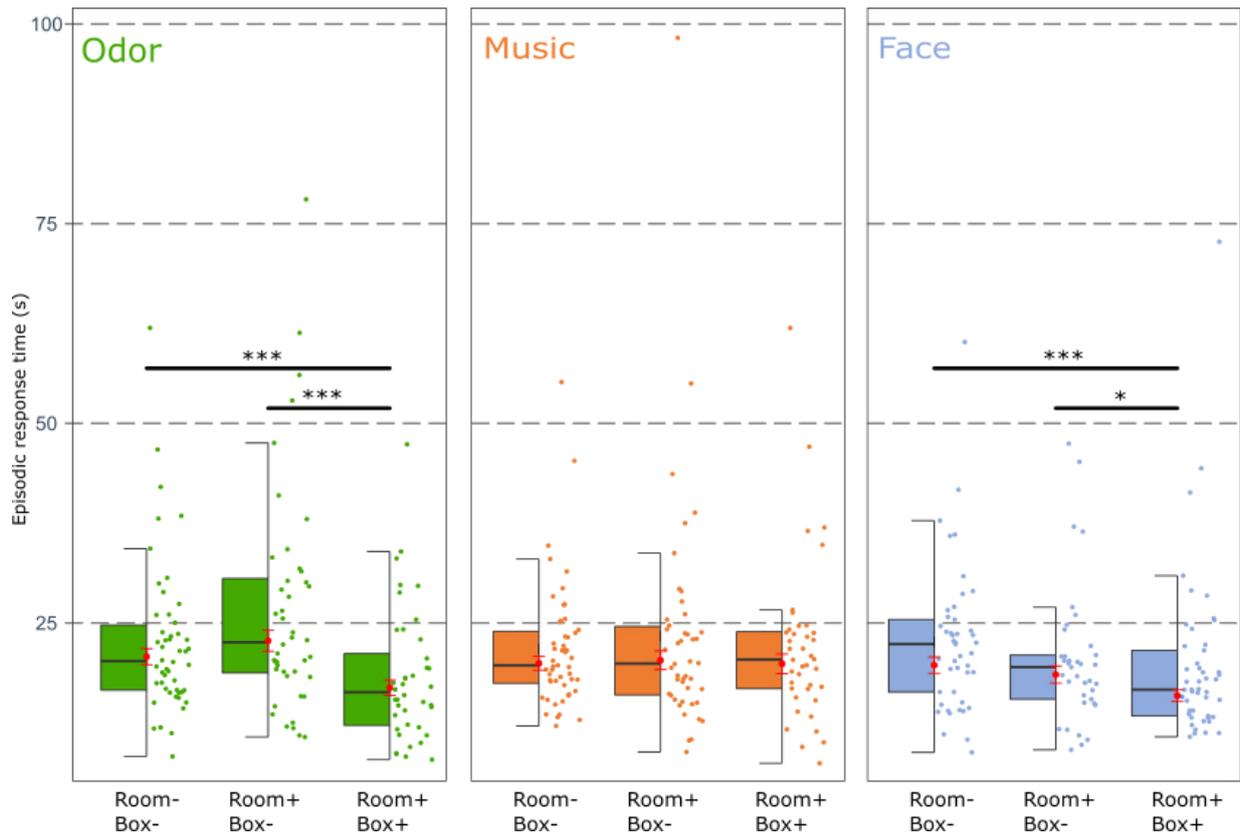


Figure 6. Episodic response times. Time (in s) to give a response (from the presentation of the cue to the click on a box) as a function of episodic response category for Odor, Music and Face. Individual raw data are represented with dots. The distribution of data is displayed with boxplot (minimum, first quartile, median, third quartile, maximum) in black. The model estimated means and their dispersion (SEM) are represented in red. * $p < .05$; *** $p < .001$.

3.4 Breathing modulations with memory performance

3.4.1 Effect of recognition responses on breathing

During recognition period, an interaction between Recognition Responses and Sensory Modality tended to be significant for *mean respiratory frequencies* ($F_{(2,98)} = 2.83$, $p = .064$) (Figure 7A). It suggested that respiratory frequency differed between Hit and CR responses ($t_{(147)} = 1.90$, $p = .059$) for odor, participants breathing more frequently for Hit than CR, but not for music and face (t 's ≤ 1.43 , p 's $\geq .15$). For *mean amplitude of inspiration*, a significant interaction between Sensory Modality and Recognition Responses ($F_{(2,98)} = 3.69$, $p < .05$) revealed that inspiration amplitude was higher during Hit than CR for music ($t_{(146)} = 2.38$, $p < .05$) and faces (t 's ≥ 2.38 , p 's $< .05$), but not odors ($t_{(146)} = 0.25$, $p = .80$) (Figure 7B). Moreover, amplitude of inspiration was always higher for odors than for faces and music whatever the recognition response (t 's ≥ 3.00 , p 's $< .01$), and was never significantly different between music and face (t 's < 0.77 , $p = 1.00$). *Mean volume of inspiration* tended to be bigger during Hit than CR ($\chi^2(1) = 3.31$, $p = .069$) responses (Figure 7C), independently of sensory modality. These

inspiration volumes were also significantly different between sensory modalities ($\chi^2(2) = 56.94$, $p < .001$) such as inspiration volumes were bigger during odor than during music and face recognition (p 's $< .001$). Such difference was not observed between music and face ($p = .86$).

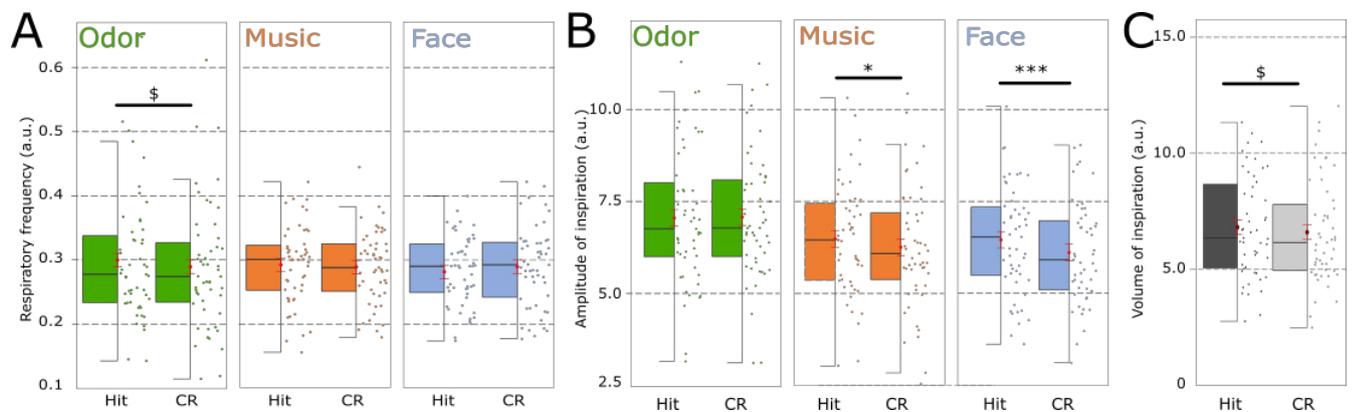


Figure 7. Breathing modulations. A. Respiratory frequency, B. Amplitude of inspiration, C. Volume of inspiration (in arbitrary units (a.u.)) during recognition period. Only differences between Hit and CR were reported. Individual raw data are represented with dots. The distribution of data is displayed with boxplot (minimum, first quartile, median, third quartile, maximum) in black. The model estimated means and their dispersion (SEM) are represented in red. * $p < .05$, *** $p < .001$, \$ $p < .07$.

3.4.2 Effect of episodic responses on breathing

During both recognition and episodic periods, no significant effects of episodic responses were observed for mean frequency, inspiration amplitude and inspiration volume of breathing. Similarly, no conclusive effects of Version were observed for any sensory modalities. See supplementary data for detailed results.

3.5 Encoding variables and their influence on memory performance

3.5.1 Encoding variables

During encoding, the Number of Presentation, *i.e.*, the amount of time each cue was explored, was of 6.32 ± 1.56 (mean \pm SD). Note that it was identical among sensory modalities because of the presence of a stimulus of each sensory modality within each box.

3.5.2 Influence of encoding variables on memory performance

The influence of encoding variables (Number of Presentation and Day) on each step of the episodic memory process is displayed in Figure 8.

3.5.2.1 From Cue to Cue +

The model used to test the influence of encoding variables on the probability to accurately recognize a target cue among all presented cues was fitted with [(0 + Number of Presentation || Participant) + (1 | Item)] as random effects (conditional $R^2 = 0.38$, marginal $R^2 = 0.33$). A

significant interaction between Sensory Modality and Number of Presentation ($\chi^2(2) = 39.99$, $p < .001$) revealed that Number of Presentation slope for odor (-0.04 ± 0.09) was significantly different from music (0.39 ± 0.12 ; $z = 5.01$, $p < .001$) and face (0.19 ± 0.10 ; $z = 5.09$, $p < .001$) ones. The slope for music tended to be different from the slope for faces ($z = 2.23$, $p = .08$). Moreover, the slope for Number of Presentation was significantly different from 0 for music ($z = 3.13$, $p < .01$), and tended to be different from 0 for faces ($z = 1.91$, $p = .06$) but not odors; $z = 0.43$, $p = 0.7$). A significant interaction between Sensory Modality and Encoding Day ($\chi^2(4) = 14.38$, $p < .01$), revealed no primacy or recency effects for odors, music and faces (z 's ≤ 2.05 , p 's > 0.12).

3.5.2.2 From Cue+ to Room+

The model used to test the influence of encoding variables on the probability to retrieve the accurate room after accurately recognizing the target cue was fitted using [(1 | Participant) +(1 | Item)] random effect structure (conditional $R^2 = 0.15$, marginal $R^2 = 0.12$). A significant interaction between Sensory Modality and Number of Presentation ($\chi^2(2) = 57.05$, $p < .001$) revealed that all slopes were significantly different between sensory modalities (z 's > 4.60 , p 's $< .001$). Moreover, Number of Presentation slope was different from 0 for faces (0.15 ± 0.05 ; $z = 2.94$, $p < .01$), but not for music (0.08 ± 0.05) and odors (0.03 ± 0.05) slopes (z 's ≤ 1.70 , p 's $> .09$).

3.5.2.3 From Room+ to Box+

The model used to test the influence of encoding variables on the probability to retrieve the accurate box after entering the accurate Room was fitted using (Encoding Day || Participants)] random effect structure (conditional $R^2 = 0.16$, marginal $R^2 = 0.14$). A significant interaction between Sensory Modality and Number of Presentation ($\chi^2(2) = 38.82$, $p < .001$) revealed that face slope (0.25 ± 0.07) was significantly different from music (0.04 ± 0.07 ; $z = 5.41$, $p < .001$), and odor (0.09 ± 0.07 ; $z = 4.39$, $p < .001$) slopes, but music and odor slopes were not significantly different ($z = 1.35$, $p = .53$). Moreover, Number of Presentation slope was different from 0 only for faces ($z = 3.37$, $p < .001$), but not for music and odors (z 's ≤ 1.24 , p 's $> .22$). A significant interaction between Sensory Modality and Encoding Day ($\chi^2(4) = 9.69$, $p < .05$) revealed no primacy or recency effects for odors, music and faces (z 's ≤ 1.96 , p 's $> .15$).

3.5.2.4 From Cue to Box+

The model used to test the influence of encoding variables on the probability to retrieve the complete episode from target and distractor cue presentation was fitted with (1 | Participant) random effect structure (BOBYQA optimizer, conditional $R^2 = 0.20$, marginal $R^2 = 0.14$). A significant interaction between Sensory Modality and Number of Presentation ($\chi^2(2) = 120.71$,

$p < .001$) revealed that face slope (0.20 ± 0.05) was significantly different from music (-0.05 ± 0.06) and odor (0.01 ± 0.06 ; z 's ≥ 7.89 , p 's $< .001$) slopes, and music and odor slopes tended to be significantly different ($z = -2.17$, $p = .090$). Moreover, Number of Presentation slope was different from 0 for face ($z = 3.71$, $p < .001$), but not for music nor odors (z 's ≤ 0.92 , p 's $> .36$).

3.5.2.5 From Cue+ to Box+

The model used to test the influence of encoding variables on the probability to retrieve the complete episode after accurate cue recognition was fitted with (1 | Participant) random effect structure (BOBYQA optimizer, conditional $R^2 = 0.20$, marginal $R^2 = 0.15$). A significant interaction between Sensory Modality and Number of Presentation ($\chi^2(2) = 119.68$, $p < .001$) revealed that all slopes were significantly different between sensory modalities (z 's > 3.84 , p 's $< .001$). Moreover, Number of Presentation slope was different from 0 for face (0.22 ± 0.06 ; $z = 3.91$, $p < .001$), but not for music (-0.05 ± 0.06) and odors (0.06 ± 0.06 ; z 's ≤ 1.04 , p 's $> .30$).

3.6 Sensory evaluations and their influence on memory performance

3.6.1 Sensory evaluations

Pleasantness evaluations (transformed with the *scale* function from base R to avoid heteroskedasticity) were analyzed with the [(0 + Sensory Modality | Participant) + (1 | Item)] random effect structure (conditional $R^2 = 0.34$, marginal $R^2 = 0.13$). A significant main effect of Sensory Modality ($F_{(2,74)} = 11.57$, $p < .001$) indicated that odors (4.42 ± 0.29) were judged as being less pleasant than music (6.39 ± 0.29) and faces (5.55 ± 0.28 ; t 's ≥ 2.76 , p 's $< .05$). No significant differences of pleasantness were observed between faces and music ($t_{(72)} = 2.07$, $p = .13$).

Emotional intensity evaluations (transformed with a *cube root* function to avoid heteroskedasticity and non-normality) were analyzed with the [(Sensory Modality | Participant) + (1 | Item)] random effect structure (conditional $R^2 = 0.24$, marginal $R^2 = 0.04$). A significant main effect of Sensory Modality ($F_{(2,82)} = 7.95$, $p < .001$) indicated that odors (5.71 ± 0.19) were judged as being more strongly emotional than music (5.05 ± 0.19) and faces (4.78 ± 0.18 ; t 's ≥ 2.67 , p 's $< .05$). No significant differences of emotional intensity were observed between faces and music ($t_{(88)} = 1.29$, $p = .60$).

Familiarity evaluations (transformed with an arcsine square root function on variable divided by 10 to avoid heteroskedasticity and non-normality) were analyzed with the [(Sensory Modality | Participant) + (1 | Item)] random effect structure (conditional $R^2 = 0.10$; marginal $R^2 = 0.04$). A significant main effect of Sensory Modality ($F_{(2,71)} = 6.66$, $p < .01$) indicated that odors (5.54 ± 0.25) were judged as being more familiar than music (4.64 ± 0.30) and faces

(4.34 ± 0.32 ; t 's ≥ 2.57 , p 's $< .05$). No significant difference of familiarity were observed between faces and music ($t_{(66)} = 0.97$, $p = 1.00$).

Complexity evaluations (transformed with an arcsine square root function on variable divided by 7.9 to avoid heteroskedasticity and non-normality) were analyzed with the [(0 + Sensory Modality | Participant) + (1 | Item)] random effect structure (conditional $R^2 = 0.12$, marginal $R^2 = 0.04$). A significant main effect of Sensory Modality ($F_{(2,75)} = 5.73$, $p < .01$), indicated that odors (5.22 ± 0.19) were judged as being more complex than music (4.24 ± 0.23 ; $t_{(84)} = 3.39$, $p < .01$). No significant differences of complexity were observed between faces (4.82 ± 0.21) and both music and odors (t 's ≤ 1.92 , p 's $\geq .18$). (see supplementary material for the interaction between version and sensory modality, which tended to be significant)

3.6.2 Influence of sensory evaluations on memory performance

The effect on individual sensory evaluations of the cues (pleasantness, emotional intensity, familiarity, complexity) on each step of the episodic memory process was represented in Figure 8. Note that, for two steps, from Room+ to Box+ and from Cue+ to Box+, the AICc backward selection presented the best model as containing only sensory modality as fixed effect and was therefore inconclusive.

3.6.2.1 From Cue to Cue +

The model used to test the influence of sensory evaluations of the cues on the probability to accurately recognize a target cue among all cues presented was fitted with [(1 | Participant) + (1 | Item)] random effect structure. The AICc backward selection presented [EI + SM * (F + C)] as the best fixed effect structure (AICc = 698.57; *BOBYQA* optimizer, conditional $R^2 = 0.46$, marginal $R^2 = 0.32$). A significant interaction between Sensory Modality and Familiarity ($\chi^2(2) = 9.57$, $p < .05$) revealed that slope for odor (0.03 ± 0.05) was significantly different from slope for music (0.40 ± 0.13 ; $z = 2.71$, $p < .05$), but not from slope for face (0.21 ± 0.08 ; $z = 1.86$, $p = .19$). Moreover, Familiarity slope was different from 0 only for music and face (z 's ≥ 2.56 , p 's $< .05$) but not odor ($z = 0.60$, $p = 1.00$). A significant interaction between Sensory Modality and Complexity ($\chi^2(2) = 6.88$, $p < .01$) revealed that the slope for face was significantly different from the slope for music ($z = 2.45$, $p < .05$). Moreover, Complexity positively influenced only face recognition (0.22 ± 0.10), as it was the only slope that tended to be different from 0 ($z = 2.19$, $p = .086$). The main effect of Emotional Intensity tended to be significant ($\chi^2(1) = 3.15$, $p = .076$), suggesting that the stronger the emotional intensity of the cue, the better the memory (0.10 ± 0.05 ; $z = 1.78$, $p = .075$).

3.6.2.2 From Cue+ to Room+

The model used to test the influence of sensory evaluations of the cues on the probability to enter the accurate room after accurate cue recognition was fitted with a [(1 | Participant) + (1 | Item)] random effect structure. The AICc backward selection presented [SM + EI + C] as the best fixed effect structure (AICc = 1365.07; *BOBYQA* optimizer; conditional $R^2 = 0.16$, marginal $R^2 = 0.12$). Significant main effects of Emotional Intensity ($\chi^2(1) = 6.16$, $p < .05$) and of Complexity ($\chi^2(1) = 4.66$, $p < .05$) revealed that the stronger the emotional intensity (0.09 ± 0.03 ; $z = 2.46$, $p < .05$) and the lower the complexity (-0.07 ± 0.03 ; $z = 2.14$, $p < .05$), the better the memory.

3.6.2.3 From Cue to Box+

The model used to test the influence of sensory evaluations of the cues on the probability to retrieve accurate episodic memory from the presentation of the cue was fitted with a (1 | Participant) random effect structure. The AICc backward selection presented [SM + EI] as the best fixed effect structure (AICc = 1252.77; conditional $R^2 = 0.19$, marginal $R^2 = 0.13$). Emotional intensity tended to have a significant effect on this probability ($\chi^2(1) = 2.98$, $p = .08$), the greater the emotional intensity, the better the memory (0.06 ± 0.04 ; $z = 1.72$, $p = .08$).

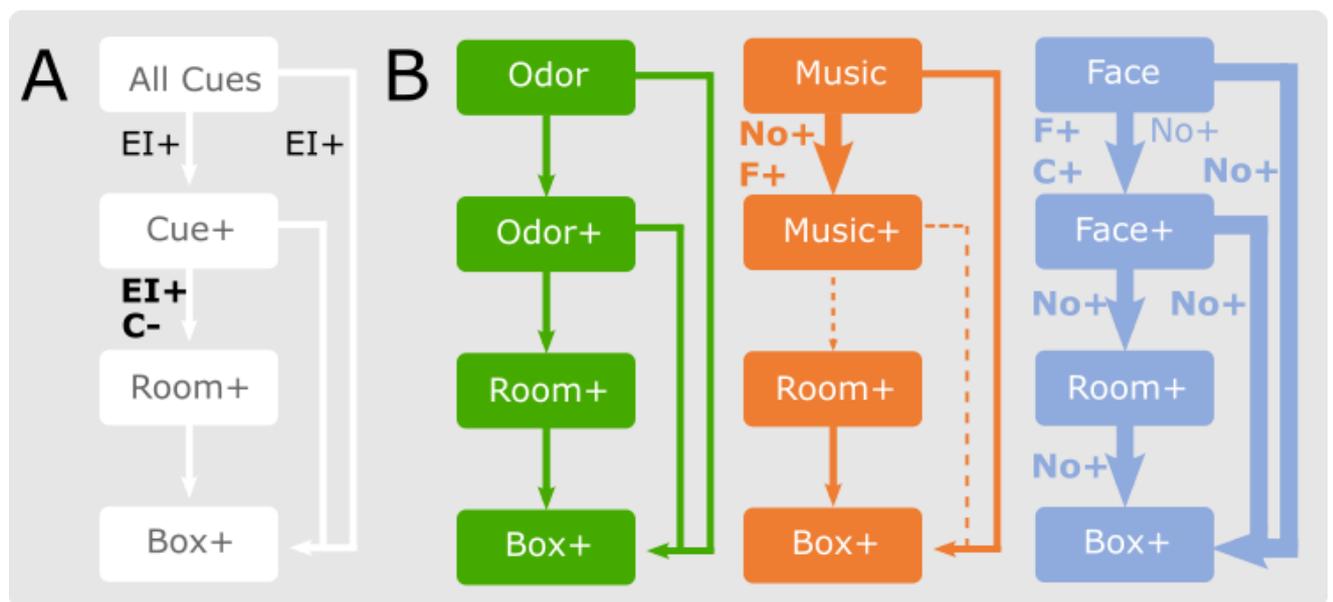


Figure 8. A. Influence of other variables on episodic memory performance. Sensory-modality aspecific, and B. Sensory-modality specific positive (+) and negative (-) influence of encoding parameters and sensory evaluations on the probability to go from one step to another. C, complexity; EI, emotional intensity; F, familiarity, No, Number of presentation. In bold, $p < .05$. In regular, $p < .09$

4 Discussion

The aims of the current study were to 1) validate our laboratory-ecological approach of EM by evaluating whether it allows for encoding of events in the lab and their retrieval after some days of delay; 2) investigate the effect of the sensory modality of the cue on the retrieval process and memories' richness, disentangling them from sensory evaluation influences; and 3) investigate respiratory behavior associated with episodic retrieval. Briefly, our results demonstrated that 1) our approach allows for the evaluation of a complex and rich EM, 2) recall patterns differed according to the sensory modality of the cue, with little effect of other sensory evaluations, and 3) breathing patterns differed according to the recognition response.

4.1 Validity of our ecological laboratory experimental approach

Our protocol used a Virtual Reality software figuring a house with three rooms deserved by a hall. This software was designed to investigate EM in a laboratory setting with ecological validity, enabling a non-explicit, incidental encoding during an immersion in a moment of life in a context with rich multisensory details. Incidental encoding avoids explicit strategies based on verbalization and imagery that might help participants to encode stimuli into memory. The software implementation also allows for motor involvement during the exploration of the event, and several (one to three) nights for memory consolidation. During the encoding phase, one room is explored at each encoding day, and in each room, three interactive boxes deliver simultaneously unfamiliar and neutral items: an odor, a music and a picture of a face. During the retrieval phase, a box placed in the hall delivers a unimodal sensory cue, either a target or a distractor. There, participants must first decide whether the cue was old (*i.e.*, part of the encoded events) or new. If the cue was considered as being old, participants had to select in which room and then in which box of the room it was delivered. The protocol thus extends our previous work that used only odor cues (Saive *et al.*, 2013, 2014b, 2015) by proposing the controlled investigation of EM evoked by unimodal cues from three sensory modalities: olfaction using smells, audition using musical excerpts and vision using face pictures. The mean number of item presentations per modality were identical, as the three stimuli were presented simultaneously each time a box was opened.

Our protocol allowed EM to be investigated sequentially: first with a recognition task, asking participants to recognize a cue among distractors; and then through a complex source memory task. This way of investigating EM has numerous advantages. Our protocol enables us to compare memory performance with chance level, validating the absence or presence of effective memory processes and strengthening the validity of the results. It allows also decomposing EM processes into sub-processes, revealing finer effects of cues or modalities. In addition to memory performance, we acquired additional measures to further evaluate the

involved processes: 1) the modulation of responses times opens a window onto task difficulty, with faster responses times suggesting easier processes. 2) The modulation of physiological responses explored through several breathing parameters allows us to enrich the characterization of memory subprocesses. And 3) the subjective feeling of the participants, who were asked to estimate whether they thought they remembered the context and place in which the memory cue was presented or not, reinforce behavioral measures.

Our results confirmed that our protocol allows for an efficient investigation of EM. When collapsing across the sensory modality of the retrieval cue, participants showed rather high performance in recognition memory (94 %) and a moderate to good performance in the whole EM retrieval (24 %) Memory performance was paralleled with response time data: correct and complete responses were faster than incorrect responses and incomplete ones, in both recognition of target cues (Hit vs. Miss) and episodic memory (Room-Box- vs. Room+Box+). Moreover, target recognition was distinguished from distractor rejection by the breathing measures: the volume of inspiration was bigger for Hit than CR. In addition, for EM, correct episodic retrieval was associated to a higher probability of subjective feeling of recollection. Correct episodic retrieval was the only condition where this probability was higher than chance and was higher than in inaccurate episodic condition (Room-Box- vs. Room+Box+).

In conclusion, our protocol allows for the investigation of efficient encoding and retrieval of sensory episodic events that are recollected with the subjective feeling of remembering. It allows the study of different steps of memory, from recognition to episodic memory and the measure of associated subjective states, in a more precise way than both the classic remember/know paradigm (Tulving, 1985) and the source paradigm (Cheke and Clayton, 2013). The use of virtual reality further ensures our protocol to be close to real life situation and to study an EM that is more ecological than what is measured in usual laboratory tasks (Schöne *et al.*, 2017).

4.2 Influence of the Sensory Modality of the cue on EM

Our study demonstrated the major influence of the sensory modality of the cue for each step of the EM process and the entire process. The sensory modality of the retrieval cue was a significant variable modulating EM processes and explaining one of many facets of memory flexibility.

4.2.1 Recognition memory

The participants were proficient in recognizing targets (Hit) with cues from all sensory modalities, and their recognition abilities were higher for both music (98%) and faces (95%) than odors (80%), confirming our hypothesis. Importantly, recognition memory was tested

using unfamiliar items, after an incident encoding and using monadic (one stimulus at a time, in contrast to a forced-choice) recognition paradigm, which are experimental conditions favoring data interpretation clarity (Cameron *et al.*, 2021). In monadic testing, contrary to forced choice testing where both a new item and an old item are presented simultaneously and where it is therefore impossible to interpret whether the decision was made on target recognition or distractor rejection, a clear analysis of detailed memory process is possible here. The use of unfamiliar items eliminates the possibility that participants were remembering the names of the items (Cessna and Frank, 2013), even if they might have used some verbal descriptions to help recognition decision. Semantic knowledge has been shown to play an important role on recognition proficiency for both odors and faces (Cornell Kärnekull *et al.*, 2015; Larsson, 1997).

The relatively good odor recognition score was in agreement with data of our preceding studies (Plailly *et al.*, 2019; Saive *et al.*, 2013, 2014b, 2015), and with data observed in other studies (*e.g.*, Larsson *et al.*, 2009; Levy, 2004). Given that the paradigms and the total number of to-be-remembered cues were different across the studies, the congruent memory performance suggests a consistency in the memory process in play. The lower recognition score for odors than other sensory cues is also in agreement with studies where recognition of odor was compared to the recognition of other stimuli such as faces (Cornell Kärnekull *et al.*, 2015), environmental sounds (Cornell Kärnekull *et al.*, 2018) and common pictures (but not abstract ones Lawless, 1978). This reduced ability to recognize unidentifiable odors can be attributed to the difficulty to verbally categorize and communicate about odors (Jraissati and Deroy, 2021; Majid and Burenhult, 2014). Indeed, a stimulus that is less easily categorized could then undergo a less deep encoding than other easily categorizable stimuli, leading to decreased recognition performance (Craik and Lockhart, 1972; Craik and Tulving, 1975). Another explanation may be that the odor perceptual system is implemented differently than other sensory systems, being for example less tied to semantization processes (Olofsson and Gottfried, 2015), or that there is poor olfactory ability linked to a poor general attention to olfaction in everyday life (McGann, 2017).

Considering music and faces, recognition memory scores were almost perfect. In light of previous findings, these results can be considered as expected. For faces, this result confirms the exceptional human ability to recognize faces, even unfamiliar ones (Sato and Yoshikawa, 2013). Memory for music was shown to be resistant, even in case of dementia, such as Alzheimer disease (Cuddy, 2018, 2018; Platel and Groussard, 2010), and can be recognized with a sense of familiarity after only two to three notes (Schulkind, 2009). Musical memory had been shown to be poorer than memory for some visual objects, such as images of object, faces, abstract art pieces, degraded visual images of scenes, and facts (Cohen *et al.*, 2009, 2011; Deffler and Halpern, 2011). Here, as there were few cues to recognize in comparison to

usual recognition study (9 musical items versus more than 60 in Cohen's study), we may thus have a ceiling effect on faces and music recognition results, not allowing to differentiate their potential differences.

4.2.2 Episodic memory

Once participants recognized the cue as being old, they had to move to the room where it was previously encountered, and then to select the box it was contained in. In order to investigate EM process, we considered either each memory step composing this process or this process as a whole. When considering EM from the cue presentation, the condition that was the closest to real-life episodic retrieval (from Cue to Room+), odor, music and faces were all able to evoke EM, with faces (44%) being more efficient than odor (19%) and music (14%). The superiority of face over odor and music cues was also observed for the subjective recollection question, revealing that participants have had stronger feeling of recollection when cued by faces than by odor and music. Suggesting the increased efficiency of faces, followed by odor cues then by music cues, the response times were influenced by episodic response accuracy for faces and odor, but not for music.

The next step (from Cue+ to Room+) which allows isolating the associative part of this memory from the recognition part, reveals that odors and faces efficiently evoked complete EM, with faces (48%) being more efficient than odors (26%), but music being unable to evoke EM (15%). Our data therefore confirmed one of the earliest results in the field of human olfactory memory, which is that the olfactory modality, in comparison with the visual modality, suffers from inferior performance during recognition, but not during associative memory (Davis, 1977). Odors are as potent as other stimuli to be associated to other objects and to be used as retrieval cues, even considering their somehow lower recognition ability (Davis, 1977; Herz, 1998; Herz and Cupchik, 1995; see Engen, 1987 for a review). This pattern of memory responses is similar to the observations made in ecological odor-evoked autobiographical memory studies suggesting that autobiographical memory seems to have a privileged relationship with the sense of smell (for reviews, Larsson & Willander, 2009; Saive *et al.*, 2014a; Hackländer *et al.*, 2019). Based on the subjective self-evaluation of memories by the participants, it is shown that, compared to the memories associated with images, sounds (*e.g.*, music, environmental sounds), words, or tactile perception, odor-related memories are more emotional (Herz, 1998; Chu & Downes, 2002). Autobiographical memories are also more emotionally charged when they are recalled using an odor cue rather than when they are recalled using the same cue but presented in another format (name, picture or sound of the odor source) (Hinton & Henley, 1993; Herz, 2004). This recall of memories evoked by smells also provides a feeling of time travel that is stronger than that evoked by images or words (Herz & Schooler, 2002). Memories associated

with smells are recalled less frequently than those associated with other sensory modalities. These memories are therefore striking; they are surprising by their little or no recalled content and by the freshness of the sensations felt (Rubin *et al.*, 1984). Finally, these memories are estimated to be older, they are mainly localized in early childhood (first decade), while memories associated with words, sounds, music or images are mainly localized in adolescence and early adulthood (second decade) (Schulkind *et al.*, 1999; Koppel & Rubin, 2016).

The separation of the associative part of the episodic memory retrieval in several steps was informative also for music memory. The first association, where participants must enter the correct room, seemed to be crucial and to be responsible of the inability of the music cue to evoke EM once recognized. While very efficiently recognized, music failed to evoke retrieval of contextual information. However, once the participants enter the correct room by chance, they were able to select the right box. The memory of the context may thus be encoded and stored but the music was not sufficient to initiate its retrieval. This weak link between music and its encoding context might have several causes. First, music might be more strongly associated to a period of life than to a specific event, and more strongly associated to social environment than to a spatial context. In their seminal work, Janata and colleagues studied the nature of autobiographical memories evoked by familiar music (Janata *et al.*, 2007), and in particular the categories of memories within the following choices: *Event, Period, Person/People, Place, None of the above*. For any given memory-evoking song, participants could select one or more of the response categories. Results showed that only one fifth of the memories were specific and related to a unique event, suggesting that music memories were more frequently associated to an ensemble of life events than to specific episodes, as in our experiment. According to the model of Conway (Conway and Pleydell-Pearce, 2000; Conway *et al.*, 2019), it is thus possible that music is more potent to trigger more strongly semantical memories, that covers a period of life or repeated events, but less prone to be associated to a precise and time-limited period of life or event.

Second, music might have elicited a strong feeling of familiarity while being almost unable to evoke recollection. This might for example explain why music recognition is spared in the case of bilateral hippocampal lesion that impairs visual memory and autobiographical memory (Esfahani-Bayerl *et al.*, 2019), the hippocampus being responsible for the encoding of the association between contextual and object elements in memory (Eichenbaum *et al.*, 2007). To test this hypothesis, it would have been interesting to identify which level of consciousness accompanied the decision of recognition in the current experiment, by adding a remember/know question to our EM paradigm, as we previously did (Saive *et al.*, 2015). We have previously demonstrated that a feeling of familiarity can be triggered by very short musical

excerpts (*i.e.*, 500-msec), suggesting the importance of potential recognition processes in music cognition (Filipic *et al.*, 2010). Interestingly, the response pattern and the sufficient amount to trigger the musical memory were similar for individuals with congenital amusia, even if their increased response times suggest difficulties in accessing the information consciously (Tillmann *et al.*, 2014).

Third, music encoding might have been disturbed by the simultaneous encoding of odor and faces. When questioned about how participants' attention was caught during encoding, participants reported that they favored odor and face encoding over music encoding. The reasons of this priority were unknown, but we could hypothesize that faces and odors are stimuli that have to be considered by humans because informing about a potential threat, while music is totally safe. It is also possible that the music used in our present study was not rewarding enough to be strongly encoded into EM. Laura Ferreri and her collaborators have indeed shown that EM performance, through a release of dopamine, was predicted by music induced reward (Ferreri and Rodriguez-Fornells, 2017; Ferreri *et al.*, 2021).

In conclusion, odors were better EM cues than music in terms of memory richness and in ease of evoking episodic memories, but they were less efficient than faces. Our hypothesis was thus only partially confirmed.

4.3 Influence of sensory evaluation of the cue on EM

Participants were asked to evaluate the pleasantness, emotional intensity, familiarity and complexity of odors, music and faces directly after retrieval. As expected based on the stimuli selection for our experiment, the cues were globally neutral, with a moderate emotional intensity, familiarity and complexity. These characteristics were chosen in order to limit other influences than the sensory modality itself. To nevertheless investigate the potential influences of individually perceived features, we linked participants' individual subjective evaluations with their memory performance. Among the measured features, the intensity of the emotion felt by the participants when perceiving the cues (even though it was not strongly modulated because of the intended neutral features of the items) influenced memory retrieval during cue recognition (from Cue to Cue+), during the retrieval of the room associated with the cue (from Cue+ to Room+) and also during the entire EM process (from Cue to Box+). The more intense was the emotion felt when perceiving the cue, the better the memory. These results are in agreement with previous findings showing that emotional intensity can influence both recognition and recollection processes (Aubé *et al.*, 2013; Saive *et al.*, 2014b; Wang *et al.*, 2021).

Beyond the role of emotional intensity of the cues on EM performance, the results revealed no effects of emotion as a function of the sensory modality of the episodic memory cues, while an

effect of valence could have been expected. It has indeed been shown that the recognition of odors, face and music was influenced by valence (Keightley *et al.*, 2011; Larsson *et al.*, 2009; Wang, 2013). This absence of effect of valence may have been caused by our stimuli selection, as all cues were selected as being as neutral as possible. These findings call for future studies investigating the specific effect of emotion on recognition and EM for the different modalities.

Familiarity of the cues favored cue recognition for music and faces, confirming the hypothesis that semantic knowledge enhances EM retrieval (Craik and Tulving, 1975). However, this effect was not observed for odors, which was a rather surprising finding as odor familiarity has been previously shown to also influence odor recognition. Familiar odors have been shown to generate a greater recollective experience in recognition memory (Larsson *et al.*, 2006), and to favor odor-evoked EM in triggering more episodic retrieval based on remembering than on the feeling of knowing (Saive *et al.*, 2015). Again, this result pattern might be explained by our choice to aim for neutral and rather unfamiliar items in the present experiment.

4.4 Interaction between EM and breathing

In human olfaction research, respiration has always been a matter of interest as odor stimuli are transported by the air to the olfactory mucosa at the top of the nostrils based on respiratory behavior. Olfactory stimulations influence breathing patterns. Unpleasant odors induce rapid and superficial breathing, while pleasant ones induce slow and deep breathing (Bensafi *et al.*, 2002; Masaoka *et al.*, 2005). Breathing parameters are therefore frequently recorded in olfactory perception and cognition studies (*e.g.*, Arshamian *et al.*, 2018; Masaoka *et al.*, 2012; Royet *et al.*, 2011; Saive *et al.*, 2014b). In the investigation of cognitive process in other sensory modalities, breathing was traditionally almost never recorded. Recently, this interest widened and an increasing number of studies reported interactions between cognitive processes and respiration patterns (Huijbers *et al.*, 2014; Iwabe *et al.*, 2014; Zelano *et al.*, 2016). For instance, Zelano *et al.* (2016) demonstrated that recognition memory accuracy was enhanced for pictures presented during inspiration rather than during expiration phases at retrieval.

In the present study, results showed a relationship between breathing patterns and recognition memory response. Importantly, for the volume of the inspiration, this effect was not modulated by the sensory modality of the cue. When target cues were accurately recognized (Hit), the volume of the inspiration increased, in comparison to when distractor cues were accurately rejected (CR). This physiological state may be related to the recollection of associated memories, as observed in our previous studies with odor cues (Saive *et al.*, 2014b). Surprisingly, while recognition of both music and faces was related to enhanced amplitude of inspiration, this effect was not significant for odors. Our hypothesis for this missing effect would

be that olfactory-associated inspiration amplitude was so high that memory modulation could not be observed. The current experiment allowed us to widen the results previously obtained in olfaction to other sensory modalities, such as audition and vision here, and to highlight the importance of investigating breathing patterns modulations with cognitive process in order to enrich this promising field of research. However, an alternative explanation would be that music and face recognition triggers the retrieval of the odor that was simultaneously presented with them, in the same box, at encoding. It is therefore important to stay cautious about the interpretation of these results, which calls for further investigation.

An olfactory specificity had been observed for the breathing pattern: odor recognition induced an increase of breathing frequency, in comparison with odor rejection, which was not observed for music and faces. When odors were recognized, participants might have wanted to extract more information about the odor stimulus by increasing the amount of sniffing. While when an odor had never been encountered and participants accurately detect odor novelty (*i.e.*, leading to correct rejection), or when a music target or a face target was presented, more information brought by sniff was not needed for the participant and breathing patterns stayed unchanged.

4.5 Conclusion

Our study demonstrated in a controlled, but ecologically relevant task the influence of the cue's sensory modality on recognition and episodic retrieval. Importantly, unfamiliar and neutral odors, music, and faces, were associated to different patterns of memory. Odors were well recognized and associated to episodic responses, musical clips were very well recognized by show limited ability to be associated to episodic dimensions, and faces were the most potent cues as they were both strongly recognized and evoked efficiently all EM dimensions. Overall, while odors were less recognized than the picture of the faces or than musical excerpts, odors demonstrate high ability for episodic retrieval, mimicking odor-evoked autobiographical memories specificities. While the present data suggest the significance of emotional intensity of the memory cue on EM, it deserved further explorations using more strongly emotional cues.

Conflict of Interest Statement

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Acknowledgments

We thank Christian Schreiber for agreeing to support our work as principal investigator and Eda Erdal for her help in data preliminary analyses. We are grateful to the technical platform *NeuroImmersion* for the development of the Episodor Virtual Reality software, and the company EmoSens® for supplying the odorants used in this study. This work was supported

by the Centre de Recherche en Neurosciences de Lyon (CRNL), the Centre National de la Recherche Scientifique (CNRS), the LABEX Cortex (NR-11-LABX-0042) of Université de Lyon within the program “Investissements d’Avenir” (ANR-11-IDEX-0007) operated by the French National Research Agency (ANR). The team *Auditory Cognition and Psychoacoustics* is part of the LabEx CeLyA (Centre Lyonnais d’Acoustique, ANR-10-LABX-60). Lucile Rey was funded by the Roudnitska Foundation.

5 References

- Abichou, K., La Corte, V., Hubert, N., Orriols, E., Gaston-Bellegarde, A., Nicolas, S., and Piolino, P. (2019). Young and Older Adults Benefit From Sleep, but Not From Active Wakefulness for Memory Consolidation of What-Where-When Naturalistic Events. *Frontiers in Aging Neuroscience* 11, 58.
- Allerborn, M., Gros, A., Messaoudi, B., Gervasoni, D., García, S., Thevenet, M., Laroche, S., Veyrac, A., and Ravel, N. (2016). A Novel Task for Studying Memory of Occasional Events in Rats. *Bio- Protocol* 6, e1740.
- Annett, J.M. (1996). Olfactory Memory: A Case Study in Cognitive Psychology. *The Journal of Psychology* 130, 309–319.
- Arshamian, A., Iravani, B., Majid, A., and Lundström, J.N. (2018). Respiration Modulates Olfactory Memory Consolidation in Humans. *J. Neurosci.* 38, 10286–10294.
- Aubé, W., Peretz, I., and Armony, J.L. (2013). The effects of emotion on memory for music and vocalisations. *Memory* 21, 981–990.
- Barr, D.J., Levy, R., Scheepers, C., and Tily, H.J. (2013). Random effects structure for confirmatory hypothesis testing: Keep it maximal. *Journal of Memory and Language* 68, 255–278.
- Bartlett, F.C. (1932). *Remembering: A study in experimental and social psychology* (New York, NY, US: Cambridge University Press).
- Bartoń, K. (2020). MuMIn: Multi-Model Inference.
- Bates, D., Mächler, M., Bolker, B., and Walker, S. (2015). Fitting Linear Mixed-Effects Models Using lme4. *Journal of Statistical Software* 67, 1–48.
- Belfi, A.M., Karlan, B., and Tranel, D. (2016). Music evokes vivid autobiographical memories. *Memory* 24, 979–989.
- Bensafi, M., Rouby, C., Farget, V., Bertrand, B., Vigouroux, M., and Holley, A. (2002). Psychophysiological correlates of affects in human olfaction. *Neurophysiol Clin* 32, 326–332.
- Bower, G.H., Monteiro, K.P., and Gilligan, S.G. (1978). Emotional mood as a context for learning and recall. *Journal of Verbal Learning and Verbal Behavior* 17, 573–585.
- Brauer, M., and Curtin, J.J. (2018). Linear mixed-effects models and the analysis of nonindependent data: A unified framework to analyze categorical and continuous independent variables that vary within-subjects and/or within-items. *Psychological Methods* 23, 389–411.
- Brédart, S., and Barsics, C. (2012). Recalling Semantic and Episodic Information From Faces and Voices: A Face Advantage. *Current Directions in Psychological Science* 21, 378–381.
- Brown, V.A. (2021). An Introduction to Linear Mixed-Effects Modeling in R. *Advances in Methods and Practices in Psychological Science* 4, 2515245920960351.
- de Bruijn, M.J., and Bender, M. (2017). Olfactory cues are more effective than visual cues in experimentally triggering autobiographical memories. *Memory* 1–12.
- Cameron, E.L., Köster, E.P., and Møller, P. (2021). Is Novelty Detection Important in Long-Term Odor Memory? *Brain Sci* 11, 1146.
- Cessna, T., and Frank, R. (2013). Does Odor Knowledge or an Odor Naming Strategy Mediate the Relationship Between Odor Naming and Recognition Memory? *Chemosensory Perception*.

- Cheke, L.G., and Clayton, N.S. (2013). Do different tests of episodic memory produce consistent results in human adults? *Learn. Mem.* 20, 491–498.
- Chu, S., and Downes, J.J. (2000). Odour-evoked autobiographical memories: psychological investigations of proustian phenomena. *Chem. Senses* 25, 111–116.
- Chu, S., and Downes, J.J. (2002). Proust nose best: Odors are better cues of autobiographical memory. *Memory & Cognition* 30, 511–518.
- Clayton, N.S., Griffiths, D.P., Emery, N.J., and Dickinson, A. (2001). Elements of episodic-like memory in animals. *Philos Trans R Soc Lond B Biol Sci* 356, 1483–1491.
- Cohen, M., Horowitz, T., and Wolfe, J. (2009). Auditory recognition memory is inferior to visual recognition memory. *Proceedings of the National Academy of Sciences of the United States of America* 106, 6008–6010.
- Cohen, M.A., Evans, K.K., Horowitz, T.S., and Wolfe, J.M. (2011). Auditory and visual memory in musicians and nonmusicians. *Psychon Bull Rev* 18, 586–591.
- Conway, M.A., and Pleydell-Pearce, C.W. (2000). The construction of autobiographical memories in the self-memory system. *Psychological Review* 107, 261–288.
- Conway, M.A., Justice, L.V., and D'Argembeau, A. (2019). The self-memory system revisited: Past, present, and future. In *The Organization and Structure of Autobiographical Memory*, (New York, NY, US: Oxford University Press), pp. 28–51.
- Cornell Kärnekull, S., Jönsson, F.U., Willander, J., Sikström, S., and Larsson, M. (2015). Long-term memory for odors: influences of familiarity and identification across 64 days. *Chem Senses* 40, 259–267.
- Cornell Karnekull, S., Arshamian, A., Nilsson, M.E., and Larsson, M. (2018). The Effect of Blindness on Long-Term Episodic Memory for Odors and Sounds. *Front. Psychol.* 9.
- Craik, F.I.M., and Lockhart, R.S. (1972). Levels of processing: A framework for memory research. *Journal of Verbal Learning and Verbal Behavior* 11, 671–684.
- Craik, F.I.M., and Tulving, E. (1975). Depth of processing and the retention of words in episodic memory. *Journal of Experimental Psychology: General* 104, 268–294.
- Cuddy, L.L. (2018). Long-Term Memory for Music. In *Springer Handbook of Systematic Musicology*, R. Bader, ed. (Berlin, Heidelberg: Springer), pp. 453–459.
- Davis, R.G. (1977). Acquisition and retention of verbal associations to olfactory and abstract visual stimuli of varying similarity. *J. Exp. Psychol.-Hum. L.* 3, 37–51.
- De Boeck, P., and Jeon, M. (2019). An Overview of Models for Response Times and Processes in Cognitive Tests. *Front. Psychol.* 0.
- DeBruine, L.M., and Barr, D.J. (2021). Understanding Mixed-Effects Models Through Data Simulation. *Advances in Methods and Practices in Psychological Science* 4, 2515245920965119.
- Deffler, S.A., and Halpern, A.R. (2011). Contextual information and memory for unfamiliar tunes in older and younger adults. *Psychology and Aging* 26, 900–904.
- Easton, A., and Eacott, M.J. (2008a). Chapter 2.3 A new working definition of episodic memory: replacing “when” with “which.” In *Handbook of Behavioral Neuroscience*, (Elsevier), pp. 185–196.

- Easton, A., and Eacott, M.J. (2008b). Chapter 2.3 A new working definition of episodic memory: replacing “when” with “which.” In *Handbook of Behavioral Neuroscience*, (Elsevier), pp. 185–196.
- Eichenbaum, H., Yonelinas, A. p., and Ranganath, C. (2007). The Medial Temporal Lobe and Recognition Memory. *Annu. Rev. Neurosci.* 30, 123–152.
- Ernst, A., Bertrand, J.M.F., Voltzenlogel, V., Souchay, C., and Moulin, C.J.A. (2021). The Proust Machine: What a Public Science Event Tells Us About Autobiographical Memory and the Five Senses. *Front. Psychol.* 11.
- Esfahani-Bayerl, N., Finke, C., Kopp, U., Moon, D.-U., and Ploner, C.J. (2019). Musical memory and hippocampus revisited: Evidence from a musical layperson with highly selective hippocampal damage. *Cortex* 119, 519–527.
- Fawcett, J.M., and Ozubko, J.D. (2016). Familiarity, but not Recollection, Supports the Between-Subject Production Effect in Recognition Memory. *Can J Exp Psychol* 70, 99–115.
- Ferreri, L., and Rodriguez-Fornells, A. (2017). Music-related reward responses predict episodic memory performance. *Experimental Brain Research*.
- Ferreri, L., Mas-Herrero, E., Cardona, G., Zatorre, R.J., Antonijoan, R.M., Valle, M., Riba, J., Ripollés, P., and Rodriguez-Fornells, A. (2021). Dopamine modulations of reward-driven music memory consolidation. *Ann N Y Acad Sci*.
- Goddard, L., Pring, L., and Felmingham, N. (2005). The effects of cue modality on the quality of personal memories retrieved. *Memory* 13, 79–86.
- Hackländer, R.P.M., Janssen, S.M.J., and Bermeitinger, C. (2018). An in-depth review of the methods, findings, and theories associated with odor-evoked autobiographical memory. *Psychonomic Bulletin & Review*.
- Harrison, X.A., Donaldson, L., Correa-Cano, M.E., Evans, J., Fisher, D.N., Goodwin, C.E.D., Robinson, B.S., Hodgson, D.J., and Inger, R. (2018). A brief introduction to mixed effects modelling and multi-model inference in ecology. *PeerJ* 6.
- Hartig, F. (2021). DHARMA: Residual Diagnostics for Hierarchical (Multi-Level / Mixed) Regression Models.
- Herz, R.S. (1998). Are odors the best cues to memory? A cross-modal comparison of associative memory stimuli. *Ann. N. Y. Acad. Sci.* 855, 670–674.
- Herz, R.S. (2004). A Naturalistic Analysis of Autobiographical Memories Triggered by Olfactory Visual and Auditory Stimuli. *Chem Senses* 29, 217–224.
- Herz, R., and Cupchik, G. (1995). The Emotional Distinctiveness of Odor-evoked Memories. *Chemical Senses* 20, 517–528.
- Herz, R., and Engen, T. (1996). Odor memory: Review and analysis. *Psychonomic Bulletin & Review* 3, 300–313.
- Herz, R.S., and Schooler, J.W. (2002). A naturalistic study of autobiographical memories evoked by olfactory and visual cues: testing the Proustian hypothesis. *American Journal of Psychology* 115, 21–32.
- Holland, A.C., and Kensinger, E.A. (2010). Emotion and Autobiographical Memory. *Phys Life Rev* 7, 88–131.
- Huijbers, W., Pennartz, C.M.A., Beldzik, E., Domagalik, A., Vinck, M., Hofman, W.F., Cabeza, R., and Daselaar, S.M. (2014). Respiration phase-locks to fast stimulus presentations:

implications for the interpretation of posterior midline “deactivations.” *Hum Brain Mapp* 35, 4932–4943.

Hutmacher, F. (2021). Do you remember? Similarities and differences between the earliest childhood memories for the five senses. *Memory* 29, 345–352.

Iwabe, T., Ozaki, I., and Hashizume, A. (2014). The respiratory cycle modulates brain potentials, sympathetic activity, and subjective pain sensation induced by noxious stimulation. *Neurosci Res* 84, 47–59.

Jaeger, T.F. (2008). Categorical data analysis: Away from ANOVAs (transformation or not) and towards logit mixed models. *Journal of Memory and Language* 59, 434–446.

Jakubowski, K., Belfi, A., and Eerola, T. (2021). Phenomenological Differences in Music- and Television-Evoked Autobiographical Memories. *Music Perception* 38, 435–455.

Janata, P., Tomic, S.T., and Rakowski, S.K. (2007). Characterisation of music-evoked autobiographical memories. *Memory* 15, 845–860.

Jiang, T., Soussignan, R., Rigaud, D., Martin, S., Royet, J.P., Brondel, L., and Schaal, B. (2008). Alliesthesia to food cues: heterogeneity across stimuli and sensory modalities. *Physiol Behav* 95, 464–470.

Jraissati, Y., and Deroy, O. (2021). Categorizing Smells: A Localist Approach. *Cognitive Science* 45, e12930.

Karlsson, K., Sikström, S., and Willander, J. (2013). The Semantic Representation of Event Information Depends on the Cue Modality: An Instance of Meaning-Based Retrieval. *PLOS ONE* 8, e73378.

Keightley, M.L., Chiew, K.S., Anderson, J.A.E., and Grady, C.L. (2011). Neural correlates of recognition memory for emotional faces and scenes. *Soc Cogn Affect Neurosci* 6, 24–37.

Kuznetsova, A., Brockhoff, P., and Christensen, R. (2015). *LmerTest: Tests in linear mixed effects models. R Package Version 2.*

Larsson, M. (1997). Semantic factors in episodic recognition of common odors in early and late adulthood: a review. *Chemical Senses* 22, 623–633.

Larsson, M., Oberg, C., and Backman, L. (2006). Recollective experience in odor recognition: influences of adult age and familiarity. *Psychol. Res.* 70, 68–75.

Larsson, M., Blåvarg, E.C., and Jönsson, F. (2009). Bad Odors Stick Better Than Good Ones: Olfactory Qualities and Odor Recognition. *Experimental Psychology* 56, 375–380.

Larsson, M., Willander, J., Karlsson, K., and Arshamian, A. (2014). Olfactory LOVER: behavioral and neural correlates of autobiographical odor memory. *Front. Psychol.* 5.

Lawless, H.T. (1978). Recognition of common odors, pictures, and simple shapes. *Perception & Psychophysics* 24, 493–495.

Lenth, R.V., Buerkner, P., Herve, M., Love, J., Riebl, H., and Singmann, H. (2020). *emmeans: Estimated Marginal Means, aka Least-Squares Means.*

Levy, D.A. (2004). Impaired odor recognition memory in patients with hippocampal lesions. *Learning & Memory* 11, 794–796.

Lo, S., and Andrews, S. (2015). To transform or not to transform: using generalized linear mixed models to analyse reaction time data. *Front. Psychol.* 0.

Lockhart, R.S., and Murdock, B.B. (1970). Memory and the theory of signal detection. *Psychol. Bull.* 74, 100–109.

- Lüdecke, D., Ben-Shachar, M., Patil, I., Waggoner, P., and Makowski, D. (2021). performance: An R Package for Assessment, Comparison and Testing of Statistical Models. *JOSS* 6, 3139.
- Ma, D.S., Correll, J., and Wittenbrink, B. (2015). The Chicago face database: A free stimulus set of faces and norming data. *Behav Res* 47, 1122–1135.
- Majid, A., and Burenhult, N. (2014). Odors are expressible in language, as long as you speak the right language. *Cognition* 130, 266–270.
- Makowski, D., Ben-Shachar, M., and Lüdecke, D. (2020). The easystats collection of R packages. (easystats).
- Maquet, P. (2001). The role of sleep in learning and memory. *Science* 294, 1048–1052.
- Masaoka, Y., Koiwa, N., and Homma, I. (2005). Inspiratory phase-locked alpha oscillation in human olfaction: source generators estimated by a dipole tracing method. *J. Physiol. (Lond.)* 566, 979–997.
- Masaoka, Y., Sugiyama, H., Katayama, A., Kashiwagi, M., and Homma, I. (2012). Slow Breathing and Emotions Associated with Odor-Induced Autobiographical Memories. *Chemical Senses* 37, 379–388.
- Matuschek, H., Kliegl, R., Vasishth, S., Baayen, H., and Bates, D. (2017). Balancing Type I error and power in linear mixed models. *Journal of Memory and Language* 94, 305–315.
- McDermott, K.B., Szpunar, K.K., and Christ, S.E. (2009). Laboratory-based and autobiographical retrieval tasks differ substantially in their neural substrates. *Neuropsychologia* 47, 2290–2298.
- McGann, J.P. (2017). Poor human olfaction is a 19th-century myth. *Science* 356, eaam7263.
- Meteyard, L., and Davies, R.A.I. (2020). Best practice guidance for linear mixed-effects models in psychological science. *Journal of Memory and Language* 112, 104092.
- Miles, A.N., and Berntsen, D. (2011). Odour-induced mental time travel into the past and future: Do odour cues retain a unique link to our distant past? *Memory* 19, 930–940.
- Olofsson, J.K., and Gottfried, J.A. (2015). The muted sense: neurocognitive limitations of olfactory language. *Trends in Cognitive Sciences* 19, 314–321.
- Plailly, J., Luangraj, N., Nicklaus, S., Issanchou, S., Royet, J.P., and Sulmont-Rosse, C. (2011). Alliesthesia is greater for odors of fatty foods than of non-fat foods. *Appetite* 57, 615–622.
- Plailly, J., Villalba, M., Vallat, R., Nicolas, A., and Ruby, P. (2019). Incorporation of fragmented visuo-olfactory episodic memory into dreams and its association with memory performance. *Sci Rep* 9, 15687.
- Platel, H., and Groussard, M. (2010). La mémoire sémantique musicale : apport des données de la neuropsychologie clinique et de la neuro-imagerie fonctionnelle. *Revue de neuropsychologie* Volume 2, 61–69.
- Royet, J.-P., Morin-Audebrand, L., Cerf-Ducastel, B., Haase, L., Issanchou, S., Murphy, C., Fonlupt, P., Sulmont-Rossé, C., and Plailly, J. (2011). True and false recognition memories of odors induce distinct neural signatures. *Front. Hum. Neurosci.* 5, 65.
- Rubin, D.C., Groth, E., and Goldsmith, D.J. (1984). Olfactory Cuing of Autobiographical Memory. *The American Journal of Psychology* 97, 493.

- Saive, A.-L., Ravel, N., Thévenet, M., Royet, J.-P., and Plailly, J. (2013). A novel experimental approach to episodic memory in humans based on the privileged access of odors to memories. *Journal of Neuroscience Methods* 213, 22–31.
- Saive, A.-L., Royet, J.-P., and Plailly, J. (2014a). A review on the neural bases of episodic odor memory: from laboratory-based to autobiographical approaches. *Frontiers in Behavioral Neuroscience* 8.
- Saive, A.-L., Royet, J.-P., Ravel, N., Thévenet, M., Garcia, S., and Plailly, J. (2014b). A unique memory process modulated by emotion underpins successful odor recognition and episodic retrieval in humans. *Frontiers in Behavioral Neuroscience* 8.
- Saive, A.-L., Royet, J.-P., Garcia, S., Thévenet, M., and Plailly, J. (2015). “What-Where-Which” Episodic Retrieval Requires Conscious Recollection and Is Promoted by Semantic Knowledge. *PLOS ONE* 10, e0143767.
- Sato, W., and Yoshikawa, S. (2013). Recognition memory for faces and scenes. *J Gen Psychol* 140, 1–15.
- Schacter, D.L., and Addis, D.R. (2007). The ghosts of past and future. *Nature* 445, 27–27.
- Schöne, B., Wessels, M., and Gruber, T. (2017). Experiences in Virtual Reality: a Window to Autobiographical Memory. *Current Psychology*.
- Schulkind, M.D. (2009). Is Memory for Music Special? *Annals of the New York Academy of Sciences* 1169, 216–224.
- Schulkind, M.D., Hennis, L.K., and Rubin, D.C. (1999). Music, emotion, and autobiographical memory: They’re playing your song. *Mem Cogn* 27, 948–955.
- Sezille, C., Messaoudi, B., Bertrand, A., Jousain, P., Thévenet, M., and Bensafi, M. (2013). A portable experimental apparatus for human olfactory fMRI experiments. *J. Neurosci. Methods* 218, 29–38.
- Singmann, H., and Kellen, D. (2019). An Introduction to Mixed Models for Experimental Psychology. In *New Methods in Cognitive Psychology*, (New York: Routledge), p. 28.
- Singmann, H., Bolker, B., Westfall, J., Aust, F., Ben-Shachar, M.S., Højsgaard, S., Fox, J., Lawrence, M.A., Mertens, U., Love, J., *et al.* (2021). *afex: Analysis of Factorial Experiments*.
- Smith, S.A. (2019). Virtual reality in episodic memory research: A review. *Psychon Bull Rev* 26, 1213–1237.
- Smith, S.M., Glenberg, A., and Bjork, R.A. (1978). Environmental context and human memory. *Memory & Cognition* 6, 342–353.
- Stalinski, S.M., and Schellenberg, E.G. (2013). Listeners remember music they like. *Journal of Experimental Psychology: Learning, Memory, and Cognition* 39, 700–716.
- Stickgold, R. (2005). Sleep-dependent memory consolidation. *Nature* 437, 1272–1278.
- Tillmann, B., Albouy, P., Caclin, A., and Bigand, E. (2014). Musical familiarity in congenital amusia: Evidence from a gating paradigm. *Cortex* 59, 84–94.
- Tolin, P., and DeGans, G.M. (1973). The effect of stimulus complexity on retrieval of information from memory. *Memory & Cognition* 1, 380–382.
- Tulving, E. (1972). Episodic and semantic memory. In *Organization of Memory*, (Oxford, England: Academic Press), pp. xiii, 423–xiii, 423.
- Tulving, E. (1983). *Elements of Episodic Memory* (Oxford: Clarendon Press).

- Tulving, E. (1985). Memory and consciousness. *Canadian Psychology/Psychologie Canadienne* 26, 1–12.
- Tulving, E., and Thomson, D.M. (1973). Encoding specificity and retrieval processes in episodic memory. *Psychological Review* 80, 352–373.
- Veyrac, A., Allerborn, M., Gros, A., Michon, F., Raguet, L., Kenney, J., Godinot, F., Thevenet, M., García, S., Messaoudi, B., *et al.* (2015). Memory of Occasional Events in Rats: Individual Episodic Memory Profiles, Flexibility, and Neural Substrate. *The Journal of Neuroscience : The Official Journal of the Society for Neuroscience* 35, 7575–7586.
- Vieillard, S., Peretz, I., Gosselin, N., Khalfa, S., Gagnon, L., and Bouchard, B. (2008). Happy, sad, scary and peaceful musical excerpts for research on emotions. *Cognition and Emotion* 22, 720–752.
- Wang, B. (2013). Facial expression influences recognition memory for faces: Robust enhancement effect of fearful expression. *Memory* 21, 301–314.
- Wang, B., Cheng, C., Jin, Z., Wu, S., and Xiang, L. (2021). The influence of negative emotional intensity on dual-processing recognition. *Biological Psychology* 161, 108083.
- Willander, J., Sikström, S., and Karlsson, K. (2015). Multimodal retrieval of autobiographical memories: sensory information contributes differently to the recollection of events. *Front. Psychol.* 6.
- Wright, D.B., and London, K. (2009). Multilevel modelling: Beyond the basic applications. *British Journal of Mathematical and Statistical Psychology* 62, 439–456.
- Wright, D.B., Horry, R., and Skagerberg, E.M. (2009). Functions for traditional and multilevel approaches to signal detection theory. *Behavior Research Methods* 41, 257–267.
- Zelano, C., Jiang, H., Zhou, G., Arora, N., Schuele, S., Rosenow, J., and Gottfried, J.A. (2016). Nasal Respiration Entrain Human Limbic Oscillations and Modulates Cognitive Function. *The Journal of Neuroscience* 36, 12448–12467.

6 Supplementary data

6.1 Debriefing questionnaires

6.1.1 Experimental procedure

Participants were asked to rate on non-graduated scales whether they paid attention to face/music/odor in their everyday life, from “none” to “a lot”. They were also told to order the three types of memory cues in terms of the attention they paid to these stimuli at the encoding phase (attention at encoding), and of the easiness to remember the room and the box when these stimuli were presented at the retrieval phase (easiness at retrieval).

6.1.2 Data analyses

The importance of each sensory modality in daily-life were analyzed. Mixed models testing the effect of daily life sensory modality importance on memory were all fitted using daily importance, sensory modality, version and the second order interaction between daily importance and sensory modality, and between daily importance and version as fixed effects. The maximal random effect models were set with daily life by-subject random slope and by-subject random intercept, including the correlation parameter. The slope or the correlation parameter were removed in case of fitting errors.

Attention at encoding and Easiness at retrieval responses involved the ranking of odor, music and face modalities (from 1 to 3). These rankings of face, music and odor were first checked with non-parametric Friedman tests, with *post-hoc* analysis with Wilcoxon signed-rank tests. As sensory ranking was not homogenous between sensory modality, it was not possible to use Sensory Modality as a fixed effect for the ranked factors, and one model was run for each of these factors. Then mixed models evaluating the effect of each of these factors on memory performance were fitted. Mixed models were fitted to predict the probability to succeed in each memory level, with as fixed effects attention at encoding or easiness at retrieval and its interaction with Version. As sensory ranking was not homogenous between sensory modality, it was not possible to put sensory modality as a fixed effect. We thus considered Sensory Modality as random effect, so maximal models were fitted with by-subject and by-sensory modality random intercepts and with the fixed effect by-subject and by-sensory modality random slopes, including correlation parameter.

6.1.3 Results

6.1.3.1 Attention at encoding

Regarding the attention at encoding: 43 participants ranked face on the first position, 7 on the second position and only one on the third position; no one ranked music on the first position,

15 participants ranked music on the second position, and 36 on the last position; 8 participants ranked odors on the first position, 37 on the second and 14 on the last position. Among ranking, the most frequent was given by 29 participants with face on the first position, then odor, and then music. The second most frequent ranking was given by 14 participants who classify face in first position, followed by music and then odor. A significant main effect of Sensory Modality ($\chi^2(2) = 60.71, p < .001$) revealed that face (1.18 ± 0.43) was ranked first and differently than odor ($2.12 \pm 0.65, p < .001$) and music ($2.71 \pm 0.46, p < .001$) cues. Odors were ranked in the second place and differently than music, which was ranked in the last position ($p < .01$).

From Cue to Cue+: The final model was fitted using by-subject and by-sensory modality random intercepts (conditional $R^2 = 0.34$, marginal $R^2 = 0.02$; BOBYQA optimizer). The statistical test revealed no effect of attention at encoding on hit probability ($\chi^2(2) = 4.92, p = 0.085$).

From Cue+ to Room+: The best model was fitted using by-subject random intercept and using attention at encoding by-sensory modality random slope, excluding the correlation parameter (conditional $R^2 = 0.07$, marginal $R^2 = 0.05$; Nelder-Mead optimizer). The analysis yielded a no significant effect of attention at encoding nor its interaction with the version of this experiment.

From Room+ to Box+: The best model was fitted using by-subject random intercept and using attention at encoding by-sensory modality and by subject random slopes, excluding correlation parameters (conditional $R^2 = 0.12$, marginal $R^2 = 0.11$; base optimizer). The analysis yielded a significant effect of attention at encoding ($\chi^2(2) = 7.64, p < .05$). Post-hoc tests revealed that rank 1 ($74.54 \pm 3.29\%$) was significantly different than rank 2 ($52.25 \pm 3.78\%$; $z = 4.3, p < .001$) and rank 3 ($38.51 \pm 4.90\%$; $z = 5.20, p < .001$), which were not different from each other ($z = -2.18, p = 0.087$).

From Cue to Box+: The best model was fitted using by-subject and by-sensory modality random intercepts and using attention at encoding by-subject random slope, including the correlation parameter (conditional $R^2 = 0.20$, marginal $R^2 = 0.14$; base optimizer). The statistical test revealed an effect of attention at encoding ranking on the probability to choose the right box after the stimulus perception ($\chi^2(2) = 24.61, p < .001$). Post-hoc tests revealed that rank 1 ($44.33 \pm 4.62\%$) was significantly different than rank 2 ($22.55 \pm 2.04\%$; $z = 4.76, p < .001$) and rank 3 ($11.33 \pm 1.99\%$; $z = 6.96, p < .001$), which were different from each other ($z = 3.68, p < .001$).

From Cue+ to Box+: The best model was fitted using by-subject and by-sensory modality random intercepts and using attention at encoding by-subject random slope, including the correlation parameter (conditional $R^2 = 0.19$, marginal $R^2 = 0.15$; base optimizer). The statistical test revealed an effect of attention at encoding ranking on the probability to choose

the right box after correctly recognizing a stimulus ($\chi^2(2) = 23.91, p < .001$). Post-hoc tests revealed that rank 1 ($49.40 \pm 4.56\%$) was significantly different than rank 2 ($26.03 \pm 2.27\%$; $z = 4.85, p < .001$) and rank 3 ($13.03 \pm 2.21\%$; $z = 7.21, p < .001$), which were different from each other ($z = 3.71, p < .001$).

6.1.3.2 Easiness at Retrieval

Regarding the easiness at retrieval: 45 participants ranked face on the first position, 2 on the second position and on the third position; one participant ranked music on the first position, 15 participants ranked music on the second position, and 33 on the last position; 3 participants ranked odors on the first position, 32 on the second position and 14 on the last position. Among ranking, the most frequent was given by 31 participants with face on the first position, then odor, and then music. The second most frequent ranking was given by 14 participants who classified face in first position, followed by music and then odor. A significant main effect of Sensory Modality ($\chi^2(2) = 62.584, p < .001$) revealed that face (1.12 ± 0.43) was ranked first and differently than odor ($2.25 \pm 0.56, p < .001$) and music ($2.63 \pm 0.56, p < .001$) cues. Odors were ranked in the second place and differently than music, which was ranked in the last position ($p < .05$).

From Cue to Cue+: The final mixed model was fitted using by-subject and by-sensory modality random intercepts (conditional $R^2 = 0.34$, marginal $R^2 = 0.03$; BOBYQA optimizer). The analysis revealed a significant effect of retrieval easiness ($\chi^2(2) = 7.73, p < .05$). Post hoc test showed that this effect was explained by the fact that there was a difference between the second ($94.98 \pm 5.24\%$) and the third rank ($90.35 \pm 5.83\%$; $z = 2.79, p < .05$), but not with the first one ($92.23 \pm 5.01\%$), which was similar to the second ($z = -1.10, p = 0.81$) and the first ($z = 0.57, p = 1.00$).

From Cue+ to Room+: The final mixed model was fitted using by-subject and by-sensory modality random intercepts and using retrieval easiness by-subject random slope, including the correlation parameter (conditional $R^2 = 0.11$, marginal $R^2 = 0.05$; base optimizer). The analysis revealed a significant effect of retrieval easiness ($\chi^2(2) = 11.95, p < .01$). Post-hoc tests revealed that rank 1 ($66.07 \pm 5.83\%$) was significantly different than rank 2 ($43.38 \pm 5.65\%$; $z = 3.60, p < .001$) and rank 3 ($45.79 \pm 5.8\%$; $z = 3.14, p < .01$), which were not significantly different from each other ($z = -0.60, p = 1.00$).

From Room+ to Box+: The final mixed model was fitted using by-subject random intercept and using retrieval easiness by-sensory modality and by subject random slopes, excluding correlation parameters (conditional $R^2 = 0.11$, marginal $R^2 = 0.10$; Nelder-Mead optimizer). The analysis revealed a significant effect of retrieval easiness ($\chi^2(2) = 8.63, p < .05$). Post-hoc tests revealed that rank 1 ($71.82 \pm 3.53\%$) was significantly different than rank 2 ($54.68 \pm 4.22\%$; z

= 3.11, $p < .01$) and rank 3 ($39.95 \pm 5.10\%$; $z = 4.37$, $p < .001$), which were not significantly different from each other ($z = 2.18$, $p = .09$).

From Cue to Box+: The final mixed model testing the effect of retrieval easiness was fitted using by-subject and by-sensory modality random intercepts and using retrieval easiness by-subject random slope, including the correlation parameter (conditional $R^2 = 0.21$, marginal $R^2 = 0.15$; Nelder-Mead optimizer). The analysis revealed a significant effect of retrieval easiness ($\chi^2(2) = 17.08$, $p < .001$). Post-hoc tests revealed that rank 1 ($44.77 \pm 4.18\%$) was significantly different than rank 2 ($20.19 \pm 2.07\%$; $z = 5.93$, $p < .001$) and rank 3 ($11.75 \pm 2.24\%$; $z = 6.93$, $p < .001$), which were different from each other ($z = 2.48$, $p < .05$).

From Cue+ to Box+: The final mixed model was fitted using by-subject and by-sensory modality random intercepts and using retrieval easiness by-subject random slope, including the correlation parameter (conditional $R^2 = 0.18$, marginal $R^2 = 0.11$; Nelder-Mead optimizer). The analysis revealed a significant effect of retrieval easiness ($\chi^2(2) = 12.53$, $p < .01$). Post-hoc tests revealed that rank 1 ($46.88 \pm 6.42\%$) was significantly different than rank 2 ($23.20 \pm 3.17\%$; $z = 3.58$, $p < .01$) and rank 3 ($15.43 \pm 3.59\%$; $z = 3.89$, $p < .001$), which were not significantly different from each other ($z = 1.71$, $p = 0.26$).

6.1.3.3 Sensory modality daily life importance

Daily life importance was changed according to the cue sensory modality. A significant main effect of Sensory Modality ($\chi^2(2) = 15.6$, $p < .001$) revealed that odors (6.98 ± 2.27) were judged to be less important than both face (7.38 ± 2.00 , $p < .01$) and music (5.68 ± 2.45 , $p < .01$) cues. Music was evaluated similarly as face cues ($p = 1.00$).

From Cue to Cue+: The final model was fitted using the maximal random effect structure (conditional $R^2 = 0.53$, marginal $R^2 = 0.12$; BOBYQA optimizer). The statistical test showed an interaction effect between, daily importance and sensory modality ($\chi^2(2) = 72.39$, $p < .001$). Post-hoc tests revealed that daily importance slope was different from 0 only for music (0.14 ± 0.05 ; $z = 2.87$, $p < .05$), but not for face (0.07 ± 0.05 ; $z = 1.51$, $p = 0.39$), nor odor (-0.07 ± 0.05 ; $z = -1.51$, $p = 0.39$). Moreover, odor slope was shown to be significantly different from face ($z = 5.89$, $p < .001$) and music ($z = 6.98$, $p < .001$) slopes, which themselves tend to be different from each other ($z = -2.32$, $p = .06$).

From Cue+ to Room+: The best model was fitted using by-subject random intercept (conditional $R^2 = 0.12$, marginal $R^2 = 0.11$; base optimizer). The statistical test revealed a principal effect of daily importance ($\chi^2(1) = 7.83$, $p < .01$), an interaction effect of daily importance by sensory modality ($\chi^2(2) = 100.98$, $p < .001$), and an interaction effect of daily importance by version ($\chi^2(2) = 4.62$, $p < .05$). Post-hoc tests revealed that daily importance

slope was different from 0 for face (0.12 ± 0.02 ; $z = 5.93$, $p < .001$) and odor (0.06 ± 0.02 ; $z = 2.45$, $p < .05$), but not for music (-0.01 ± 0.02 ; $z = -0.54$, $p = 1.00$). Moreover, face slope was significantly different from music ($z = 9.63$, $p < .0001$) and odor ($z = 3.91$, $p < .001$) slopes, which themselves were significantly different from each other ($z = -4.47$, $p < .001$). Moreover, the global slope for version 2 (0.07 ± 0.02) was significantly different from 0 ($z = 3.25$, $p < .01$) and from the first version slope ($z = 2.19$, $p < .05$). The slope for the first version tended also to be different from zero ($z = 2.07$, $p = 0.078$).

From Room+ to Box+: The maximal model was used to predict the probability to choose the correct box while being in the correct room (conditional $R^2 = 0.14$, marginal $R^2 = 0.06$; base optimizer). The statistical test revealed an interaction effect of daily importance by sensory modality ($\chi^2(2) = 26.77$, $p < .001$). Post-hoc tests revealed that daily importance slope was different from 0 for face (0.09 ± 0.03 ; $z = 3.11$, $p < .01$), but not for music (-0.01 ± 0.03 ; $z = -0.16$, $p = 1.00$), nor odor (0.04 ± 0.03 ; $z = 1.07$, $p = .86$). Moreover, face slope was significantly different from music ($z = 4.96$, $p < .001$) and odor ($z = 2.59$, $p < 0.05$) slopes, which themselves were not significantly different from each other ($z = -1.82$, $p = 0.21$).

From Cue to Box+: The best model was fitted using by-subject random intercept (conditional $R^2 = 0.18$, marginal $R^2 = 0.13$; BOBYQA optimizer). The statistical test revealed a principal effect of daily importance ($\chi^2(1) = 6.69$, $p < .01$) and an interaction effect of daily importance by sensory modality ($\chi^2(2) = 106.81$, $p < .001$). Post-hoc tests revealed that daily importance slope was different from 0 only for face (0.14 ± 0.02 ; $z = 6.16$, $p < .001$), but not for music (0.00 ± 0.02 ; $z = -0.14$, $p = 1.00$), nor odor (0.04 ± 0.03 ; $z = 1.54$, $p = 0.37$). Moreover, face slope was shown to be significantly different from music ($z = 9.37$, $p < .001$) and odor ($z = 6.11$, $p < .001$) slopes, which themselves were significantly different from each other ($z = -2.55$, $p < .05$).

From Cue+ to Box+: The best model was fitted using a daily by-subject random intercept including the correlation parameter (conditional $R^2 = 0.20$, marginal $R^2 = 0.13$). The statistical test revealed a principal effect of daily importance ($\chi^2(1) = 6.50$, $p < .05$) and an interaction effect of daily importance by sensory modality ($\chi^2(2) = 99.70$, $p < .001$). Post-hoc tests revealed that daily importance slope was different from 0 only for face (0.20 ± 0.04 ; $z = 5.37$, $p < .001$), but not for music (-0.02 ± 0.04 ; $z = -0.54$, $p = 1.00$), nor odor (0.09 ± 0.04 ; $z = 2.10$, $p = 0.11$). Moreover, face slope was shown to be significantly different from music ($z = 9.56$, $p < .001$) and odor ($z = 4.38$, $p < .001$) slopes, which themselves were significantly different from each other ($z = -4.182$, $p < .001$).

6.2 Supplementary breathing data

6.2.1 Breathing results

6.2.1.1 Effect of odor recognition responses on breathing

6.2.1.2 Effect of the subsequent episodic response on the breathing during recognition period

Mean respiratory frequency: A 2 (Version) x 2 (Recollection score: Room-Box-, Room+Box-, Room+Box+) rmANOVA revealed no significant difference among effects for faces. The same analysis on music and a non-parametric Friedman test investigating the recollection score effect on odor mean frequency revealed no significant difference among effects. However, a Kruskal-Wallis rank sum test done on odor mean frequency testing the version effect revealed significant differences among versions ($\chi^2(1) = 4.57, p < .05$), version 1 (0.30 ± 0.11) yielding to breathing with higher respiratory frequency during odor perception than version 2 (0.29 ± 0.10).

Inspiratory amplitude: A 2 (version) x 2 (recollection score) rmANOVA revealed no significant difference among effects for face. The same analysis for music revealed a significant interaction effect between version and episodic score ($F_{(2,66)} = 8.58, p < .001$). Post-hoc tests revealed that inspiration amplitude did not differ between version for each music score ($p > .15$), but that amplitude was lower after Room+Box- (5.85 ± 0.36) than after Room+Box+ (6.81 ± 0.36 ; $t_{(66)} = -3.93, p < .001$) and Room-Box- (6.49 ± 0.36 ; $t_{(66)} = 2.62, p < .05$) for version 1 and no for version 2 ($p > .17$). A non-parametric Friedman test investigating the recollection score effect on odor inspiration amplitude revealed no significant difference among effects ($\chi^2_{(2)} = 2.24, p = .33$). However, a Kruskal-Wallis rank sum test done on odor mean frequency testing the version effect revealed significant differences among versions ($\chi^2_{(1)} = 14.24, p < .001$), version 1 (7.54 ± 2.27) yielding more frequent breathing during odor perception as version 2 (6.51 ± 1.07).

Inspiratory volume: A 2 (version) x 2 (recollection score) rmANOVA revealed no significant difference among effects for faces. The same analysis on music was run and revealed a significant interaction effect between version and episodic score ($F_{(2,66)} = 4.52, p < .05$). *Post-hoc* tests revealed that inspiration volume did not differ between version for each music score ($p > .25$), but that volume was lower after Room+Box- (4.89 ± 0.43) than after Room+Box+ (5.77 ± 0.43 ; $t_{(66)} = 2.89, p < .05$) for version 1 and no for version 2 ($p = 1.0$). A non-parametric Friedman test investigating the recollection score effect on odor inspiration volume revealed no significant difference among effects ($\chi^2_{(2)} = 2.18, p = .34$). However, a Kruskal-Wallis rank sum test done on odor mean volume testing the version effect revealed no significant differences among versions ($\chi^2_{(1)} = 0.11, p = .74$).

6.2.1.3 Effect of the episodic response on the breathing during episodic period

Respiratory frequency: 2 (version) x 2 (recollection score) rmANOVAs were conducted for faces, music and odors and revealed no significant difference among effects.

Inspiratory amplitude: A non-parametric Friedman test investigating the recollection score effect for face revealed no significant difference among effects ($\chi^2(2) = 0.06, p = .97$). However, a Kruskal-Wallis rank sum test done on odor mean amplitude testing the version effect revealed a tendency showing differences among versions ($\chi^2(1) = 2.98, p = .084$). A 2 (version) x 2 (recollection score) rmANOVA was conducted for music and revealed no significant difference among effects. The same analysis was run for odors and revealed a significant interaction effect of Version ($F_{(1,31)} = 5.69, p < .05$), inspiration amplitude in version 1 (7.47 ± 0.29) being higher than for version 2 (6.48 ± 0.29).

Inspiratory volume: A 2 (version) x 2 (recollection score) rmANOVA was conducted on face inspiration volume and revealed no significant difference among effects. The same analysis on music and odor revealed no significant difference among effects.

6.3 Supplementary sensory evaluations data

6.3.1.1 Complexity effect on sensory modality

Complexity evaluations (transformed with an arcsine square root function on variable divided by 7.9 to avoid heteroskedasticity and non-normality) were analyzed with the [(0 + Sensory Modality | Participant) + (1 | Item)] random effect structure (conditional $R^2 = 0.12$, marginal $R^2 = 0.04$). A significant interaction between Sensory Modality and Version ($F_{(2,75)} = 2.78, p = .072$), indicated that complexity was higher for face than for music in Version 1 ($p < .05$), but not in Version 2 ($p = .10$).

Study 2

Episodic memory evoked by odors compared to musical excerpts and faces: Modality-specific effect of wanting, but modality-unspecific effect of valence

Lucile Rey^{1-2 *}, Clément Désoche³, Marc Thévenet¹⁻², Samuel Garcia¹⁻², Barbara Tillmann^{2*}, Jane Plailly^{1*}

¹ CMO team, Lyon Neuroscience Research Center, CNRS UMR 5292 - INSERM U1028 - University Lyon1, Lyon, France

² CAP team, Lyon Neuroscience Research Center, CNRS UMR 5292 - INSERM U1028 - University Lyon1, Lyon, France

³ IMPACT team, Lyon Neuroscience Research Center, CNRS UMR 5292 - INSERM U1028 - University Lyon1, Lyon, France

⁴Hospices Civils de Lyon, Mouvement et Handicap, Neuro-Immersion, Lyon, France

*The last two authors contributed equally to this work

Running title: Wanting favors odor-evoked episodic memory

***Correspondence:**

Lucile Rey

Centre de Recherche en Neurosciences de Lyon, INSERM U1028 - CNRS 5292 - UCBL - UJM

Centre Hospitalier Le Vinatier, Bâtiment 462 - Neurocampus Michel Jouvét

95 boulevard Pinel 69675 Bron Cedex FRANCE

e-mail: lucilerey.pro@gmail.com

Abstract

The role of emotion on memory is well-known. Emotions influence both the content of and the access to personal event memories (Burke *et al.*, 1992; Sheldon *et al.*, 2020). In particular, odors and music are shown to evoke particularly emotional episodic memories. However, the effect of emotion on episodic memory may depend on sensory modality of the memory-triggering cue. The aim of this study was to unravel the effect of emotion and sensory modality as well as its potential interaction on performance in an episodic memory task that is close to ecological conditions, but ran in a laboratory setting. Using a non-immersive virtual reality device presenting a three-room house, participants freely and incidentally explored three unique rich episodes over three consecutive days. Episodes were constructed around the three dimensions used to characterize episodic memory: What (odor, music, face), Where (the rooms: bedroom, living room, office), and in *Which context* (the periods of the day: daytime, nighttime or twilight). The retrieval was tested on the fourth day. Participants were told to recognize encoded stimuli among distractors and to select both the room and the period in which they thought they encountered the stimulus at encoding. This paradigm and its analyses allowed us to study the episodic memory processes as a whole as well as step by step, and the interaction between the sensory modality of the cue and the emotional content of the cue, as evaluated individually, on episodic memory performance. Results demonstrated that the emotion of the cue had a global effect on almost all memory steps, positively and negatively valenced emotional cues improving memory performance in comparison to neutral cues. The sensory modality of the cues also affected memory performance, and the influence of emotional content on memory process differed with sensory modalities. Odors were shown to be potent cues to evoke strong episodic memories, and the whole episodic retrieval was influenced by the motivation (wanting) towards the odor. Music was highly recognized and this recognition was favored by emotion, but they were unable to evoke episodic memory. Faces were able to evoke only one episodic dimension retrieval, and this retrieval was moderately enhanced with faces emotion. Results are discussed regarding the influence of emotion and wanting in recognition and episodic retrieval.

Keywords: Episodic memory; Recognition memory; Valence; Emotional intensity; Wanting; Odors; Music; Faces; Virtual Reality.

1 Introduction

Episodic memory (EM) can be defined as the conscious re-experience of unique personal events from the past (Tulving, 1972), or the memory of *what* happened, *where* and in *which context* (Cheke and Clayton, 2013; Easton and Eacott, 2008a; Tulving, 1983). Episodic memories are not the exact copy of the initial event, they are different each time they are retrieved (Doerksen and Shimamura, 2001; Ernst *et al.*, 2021; Hutmacher, 2021; Kensinger and Corkin, 2003; Rey *et al.*, a; Saive *et al.*, 2014a). For example, emotion is a highly powerful memory cue that influences episodic memory retrieval (Kaplan *et al.*, 2016; Levine and Pizarro, 2004). The cue's emotion influences the valence (Schulkind and Woldorf, 2005) and the amount of details of the memories. Valence can change the type of memories recalled and high arousing cues lead to more details (Sheldon and Donahue, 2017; Sheldon *et al.*, 2020). Interestingly, highly emotionally intense cues improve the memory of the central item of the event while impairing the memory of the context (Bisby *et al.*, 2016; Burke *et al.*, 1992; Mao *et al.*, 2015; Palombo *et al.*, 2018).

The effect of emotion on recognition and episodic memory has been observed in several sensory modalities. The recognition of odors was influenced by their valence (Larsson *et al.*, 2009) and emotional intensity (Saive *et al.*, 2014a). The recognition of music is favored by their liking (Stalinski and Schellenberg, 2013), reward value (Ferreri and Rodriguez-Fornells, 2017; Ferreri *et al.*, 2021), emotional intensity (Aubé *et al.*, 2013; Eschrich *et al.*, 2008), and arousal (Alonso *et al.*, 2015; Ferreri and Rodriguez-Fornells, 2017; Ferreri *et al.*, 2021). Moreover, both emotional intensity and positive valence were shown to be potent dimensions related to autobiographical salience of musical pieces (Salakka *et al.*, 2021). Emotional music evokes more detailed AM than less emotional music (Janata *et al.*, 2007). Regarding faces, the results seem contradictory. Negative faces have been reported to be better recognized than neutral ones (Keightley *et al.*, 2011; Wang, 2013). However, in other studies, face recognition was not influenced by valence (Johansson *et al.*, 2004) and recognition was better for neutral faces (Cortes *et al.*, 2017). Similarly, in remember/know paradigm, negative emotions of faces are sometimes associated to more (Johansson *et al.*, 2004; Wang, 2013) or less (Cortes *et al.*, 2017) remember responses than neutral faces. These discrepancies might be linked to the use of *a priori* categorized emotional stimuli rather than with stimuli that are based on participant's subjective feelings relative to the displayed emotion. Indeed, both odors and music perception may be subject to high inter-individual variability (Mantel *et al.*, 2019; Mas-Herrero *et al.*, 2013, 2014). Taking in consideration the subjective evaluation seems thus more appropriate than using *a priori* categorized emotional stimuli.

The sensory modality of the cue triggering episodic retrieval is suggested to be another parameter influencing memories' content and phenomenological features of this remembering, and in particular the emotional dimension. These observations were made by directly comparing episodic memory triggered by cues from different sensory modalities. It has been suggested that odor-evoked memories are different from memories evoked by other sensory cues, even though it is the same episodic event that is retrieved (Annett, 1996; Hackländer *et al.*, 2018; Herz and Engen, 1996; Larsson *et al.*, 2014; Saive *et al.*, 2014b). The emotional power of odors and the proximity of its neural correlates with the amygdala led to the hypothesis that the specificity of odor-evoked EM is related to the link between odors and emotional processes (Larsson *et al.*, 2014). Research shown that odors evoke less memories than other sensory modality cues, but once autobiographical memories are remembered, these memories are more emotional than when evoked with cues from other sensory modalities, such as pictures, sounds and labels. Similarly, the memory of pictures associated with odors is more emotionally loaded than the memory of pictures associated with verbal or tactile cues (Herz, 1998; Herz and Cupchik, 1995). Music-evoked autobiographical memories are also reported to be specifically associated with strong emotions, and these emotions are more intense than those related to autobiographical memories evoked by faces or by a TV show (Belfi *et al.*, 2016; Jakubowski *et al.*, 2021a). However, the direct comparisons of episodic memory triggered by emotional cues from different sensory modalities are currently still understudied and need to be further investigated.

In a previous study (Rey *et al.*, a), we investigated the effect of the sensory modality of neutral cues on EM retrieval. For that purpose, we developed a new experimental protocol, using virtual reality to favor EM encoding and retrieval (Smith, 2019), allowing the direct comparison of EM processes evoked by three different stimuli: odors, musical excerpts and faces. This study revealed the following main findings. While all cues were significantly recognized, musical pieces and faces were better recognized than odors. Moreover, odors and faces were more potent to allow the retrieval of the context of previously encountered events than were music stimuli. Finally, exploratory analyses suggested that, when considering all cues together, the retrieval of episodic information was favored by the emotional intensity of the cue, as evaluated individually by each participant. These findings provide evidence that EM is influenced by the sensory modality of the retrieval cue and suggest that this effect might be moderated by the emotion conveyed by the cue. Based on this study and on previous findings showing a unique link between odor memory and emotion (Herz and Cupchik, 1995; Larsson *et al.*, 2014; Saive *et al.*, 2014a), we hypothesized that this effect of emotion on EM retrieval may differ as a function of the cue's sensory modality.

The present study aimed to investigate the hypothesis that the effect of emotion on episodic memory depends on the sensory modality of the cue. This hypothesis was tested with an ecological laboratory protocol in which participants discovered events in semi-immersive virtual reality. During a first-person free exploration, participants visited a three-room (bedroom, living-room, office) house where specific devices diffusing an odor, a musical excerpt, or a picture of a virtual face were disposed. Each room contained three devices (one for each modality), all diffusing a stimulus of the same emotional valence. Three emotional categories of stimuli differing in valence were used: unpleasant (labelled as negative), neutral and pleasant (labelled as positive). Aiming to restrict potential influences of semantic processes, odors, music, and faces were selected to be as weakly identifiable as possible. As for most real-life events, the encoding was not explicit, but participants freely explored the house on three consecutive days, without the memory task being announced. The house was identical across days, but the period of the day differed, being daytime, nighttime or twilight, so did the displayed emotion in each room and all stimuli. On the fourth day, EM retrieval was tested with cues being odors, music or faces, either previously presented at encoding or not, and participants performed a series of tasks. EM was investigated with the evaluation of the probability to recognize the memory cue among distractors, and in case of recognition, to remember the period of the day and the room in which the cue was initially encoded. At the end of the testing session, participants evaluated the sensory features of the memory cues with four subjective scales, pleasantness, emotional intensity, familiarity and wanting. Based on previous findings observed for EM of odor, music and face cues, our hypothesis was that episodic memory would be favored by the emotional content of the cues for the three sensory modalities, as evaluated from the stimulus' subjective evaluations regarding emotional dimensions, but that this effect would be stronger for odors than for music and faces.

2 Materials and Methods

2.1 Ethics Statement

All participants provided written informed consent prior to participation. The protocol was approved by the national Institutional Review Board, according to French regulations for biomedical experiments with healthy volunteers [Ethical Committee of CPP IDF8 (April 25,2017), ID RCB: 2016-A01931-50].

2.2 Participants

Thirty-three healthy participants [21 women; aged 23.9 ± 4.8 (mean \pm standard deviation)] consented to participate in this experiment and received 40 Euros in compensation. All participants reported normal senses of smell and no auditory or visual impairments. They were recruited through campus electronic mails and posters, and social networks. Given the implicit

nature of the experiment, they were told that the study aimed to investigate the perception of various environments involving pictures, music, and odors. Seven participants reported to have on average 4.6 (± 2.7) years of formal musical training on an instrument (minimum of 6 months, maximum of 8 years).

2.3 Stimuli and materials

2.3.1 Multidimensional episodes

A software called EpisOdor-Version2 was developed using Unity 5.6.5 (Unity Technologies, USA) by a local technical platform *NeuroImmersion* (Cesame Institute, CRNL). EpisOdor-Version2 allowed the presentation of a virtual 3D house in a first-person view (non-immersive virtual reality) that can be actively explored using a trackball (Kensington, Redwood Shores, CA, USA (Figure 1A)). This house was presented in three conditions created with variation of light intensity and color: day, night, and twilight (defining the *Which context* component). The house was composed of three rooms (defining the *Where* component), a bedroom, an office and a living room, connected to a corridor by closable doors. Each of these rooms was defined by its furniture (*e.g.*, chair, bed, table, piano) and decorative elements (*e.g.*, painting, carpet). In addition, three clickable objects, a perfume diffuser, a radio set and a photo frame, were placed in each room to deliver respectively an odor, a music and a picture of a face (defining the *What* component) (Figure 1B). They were highlighted by an arrow that appeared when participants were nearby. These objects were relatively distantly located from each other within a given room and were arranged differently between each room. To avoid any emotional contagion between stimuli in the same room, each room was assigned to one emotion, all rooms presented in the same context being associated to different emotions. This assignment was set randomly and varied with the house context (period of the day), so that each room was associated with all emotions.

2.3.2 Stimuli and apparatus

To favor the comparability between the stimuli of different sensory modalities, eighteen stimuli of each sensory modality (odor, music, face) and emotion category (negative, neutral, positive) were selected based on a behavioral pre-test in which 12 participants evaluated the items on subjective scales from 0 (minimal) to 10 (maximal) for valence, emotional intensity, low familiarity and complexity. The 18 stimuli of each sensory modality were pseudo-randomly assigned as target or distractor items for each participant, so that 9 were target cues and 9 were distractor cues. Each set of cues (target and distractor) consisted of three stimuli per emotion category for each sensory modality.

2.3.2.1 Odorants

Odorants consisted of essential oils, single or mixtures of monomolecular chemical compounds, and fragrances were used. Negative odorants were composed of Carrot (Givaudan-Roure[®], Vernier, Suisse), Musk (Givaudan-Roure), Pastry (Perlarom 93, Grasse, France), Stemone (Créations aromatiques), Vetiver (Davenne, Montfavet, France), and Yeast (Givaudan). Neutral odorants were composed of Basil (Créations aromatiques, Neuilly-Sur-Seine, France), Black Olive (Meilleur du Chef[®], Bassussarry, France), Cis-3-hexenyl salicylate (Créations aromatiques), Dill (Pharmacie Croix Blanche, Dijon, France), Honey (Givaudan-Roure), and Ylang Ylang (Givaudan). Positive odorants were composed of Bien-Être (EmoSens), Blackberry (Givaudan[®], Vernier, Suisse), Cosy (EmoSens), Lovely Ion (EmoSens[®], Lyon, France), Osmose (EmoSens), and Tomato (Givaudan-Roure, Vernier, Suisse). The undiluted odorants were placed in a 10 ml U-shaped Pyrex[®] tubes (VS Technologies, Saint-Priest, France) filled with Pebax[®] microporous substances (Arkema, Colombes, France).

The odorants were presented with a 20-channel computer-controlled olfactometer adapted from (Sezille *et al.*, 2013). This odor diffusion system was developed to synchronize odorous stimuli with breathing. The participants' nasal respiratory signals were acquired using a nasal canula and were used to trigger the odor stimulation through an airflow sensor. During odor stimulation, the olfactometer waited for the participants' subsequent expiration, allowing the odor to be perceived at the beginning of the following inspiration. When this expiration was detected, an unodorized airflow was sent to one of the U-shaped odorous tubes. Odorized airflow and air carrier were sent to and mixed in a homemade mixing head made of polytetrafluoroethylene. It was connected to the nostrils through two Teflon tubes, fixed to the nasal canula, opening out under the nostrils. The olfactometer airflow rate was set between 3.0 and 4.5 l/min depending on the odorant physical intensity. The odorants were delivered over 6 s. The olfactometer was controlled by an in-house LabView software (National Instruments, Austin, TX, USA) and interacted with EpisOdor-Version2 software to send odor after a click on the perfume diffuser.

2.3.2.2 Musical Pieces

Musical clips were selected from the materials used in Vieillard *et al.* (2008) (Copyright, Bernard Bouchard, 1998) (http://www.peretzlab.ca/knowledge_transfer/) and from our previous experiment (Rey *et al.*, a). They were modified in MIDI (Digital Performer[®], MOTU, Cambridge, USA) aiming to create negative, neutral and positive musical excerpts. The modifications included changes of tempo, mode (major/minor), and/or a few notes, depending on the clip and the targeted emotion. The musical clips were played with an acoustic piano

timbre (Cubase[®], Steinberg Media Technologies, Hamburg Germany) and were presented with EpisOdor-Version2 at a comfortable loudness level with headphones. Their average duration was 7.11 ± 2.14 s.

2.3.2.3 Faces

Faces were selected from the Compound Facial Expressions of Emotions Database (Du *et al.*, 2014) (http://cbcs1.ece.ohio-state.edu/dbform_compound.html). They were composed of nine women and nine men faces from Caucasian origin (Figure 1C). These faces were virtualized with CrazyTalk^{®8} software (Reallusion Inc., California, U.S.A.), by selecting the head shape that best matched an original photography of a face depending on its origin, and then by selecting more finely the contour and textures that best matched the face. The same non-gendered hairstyle and outfit were chosen for each face. Pictures of this virtualized faces were then turned into black and white, presented on a white background, and their relative position and dimensions were similar (spots, freckle; Photoshop[®], Adobe, Dublin, Ireland). The 2863 x 1718 pixel pictures were presented in jpeg format inside the picture frame, on the center of the screen, with EpisOdor-Version2. The duration of the presentation of each face was of 6 s.

2.4 Experimental procedure

The experiment was divided into four sessions, one session per day on consecutive days (Figure 1D). The first three days were devoted to the encoding phase, and the last day was divided into three sessions: retrieval test, questionnaire, and stimuli sensory evaluation. A night of sleep followed each of the encoding sessions to promote consolidation and to reduce interference (Maquet, 2001; Stickgold, 2005). Participants completed the four sessions at the same time of the day to limit the differential influences of internal states (hunger, satiety) on olfactory and cognitive processes between sessions (Jiang *et al.*, 2008; Plailly *et al.*, 2011). The participants were requested to breathe through their nose as normally as possible without consciously modifying their respiration.

2.4.1 Encoding ¹¹

Each encoding session consisted of at least 10 min during which participants freely explored the house presented either at daytime, nighttime or twilight on each of the three days. The order of these three time periods (*i.e.*, referred to as Periods thereafter) was counterbalanced across participants. Participants were encouraged to pay attention to each detail of the rooms and were told to click at least five times on each clickable item to be sure they had sufficiently explored each stimulus. The fifth click was indicated to the participant with a change of the

¹¹ Note: due to software issues, some clicks on the perfume diffuser were not considered by the olfactometer. Some participants have thus explored odors less than five times.

color of the arrow over the interactive device, from green to red. No instructions for memorization were given, but participants were informed that they would be questioned about their perception of the episodes on the fourth day. This procedure aimed to ensure free encoding like real-life situations.

2.4.2 Retrieval

The retrieval test was displayed with an in-house LabView software and was composed of 54 trials (Figure 1E). The session started after 10 s of pause. Each trial started by 2 s of pause, then a unimodal cue (Odor, Music or Face) was presented. Half of the memory cues had never been encountered during the encoding phase (27 distractors), and the others had been explored during the three first days (27 targets). The cues were presented pseudo-randomly with the constraint to avoid the presentation of two odorants, or three music or faces in a row. This avoided olfactory saturation, and any habituation to the stimulus type. Participants were first asked whether they recognized the cue as having been presented at the encoding phase or not (Recognition response: "*Do you recognize this stimulus? (Y/N)*"). The answer was given with left (Yes) or right (No) click on the trackball. In the case of a "No" response, the trial ended. Therefore, the cue diffusion stopped, and a black screen with the white inscription "Click whenever you want" was displayed. After a click, another trial started. In the case of a "Yes" response, the participants were asked to evaluate their subjective recollection of the event evoked by the cue ("*Do you remember the context? (Y/N)*"), allowing metacognitive knowledge evaluation. In case of a "No" response, a black screen with a white cross in the middle appears for 5 s. In case of a "Yes" response, a black screen with the white inscription "Immerse yourself into this memory" was displayed. Then, regardless of the response, a white screen with 6 images arranged in 2 rows x 3 columns appeared. The three images in the upper side of the screen represented three pictures of the house taken from an aerial view, with respectively the day, twilight, and night contexts. These photos were placed above the question "In which context did you discovered it?". The three images in the lower side represented three pictures of the rooms that were turned into black and white, with respectively the office, the living room, and the bedroom. These photos were placed above the question "In which room did you discovered it?". Participants had to click on one context and one room picture in any order to end the trial (Episodic response). Once the response was given, participants were given an additional two seconds to eventually change their selection.

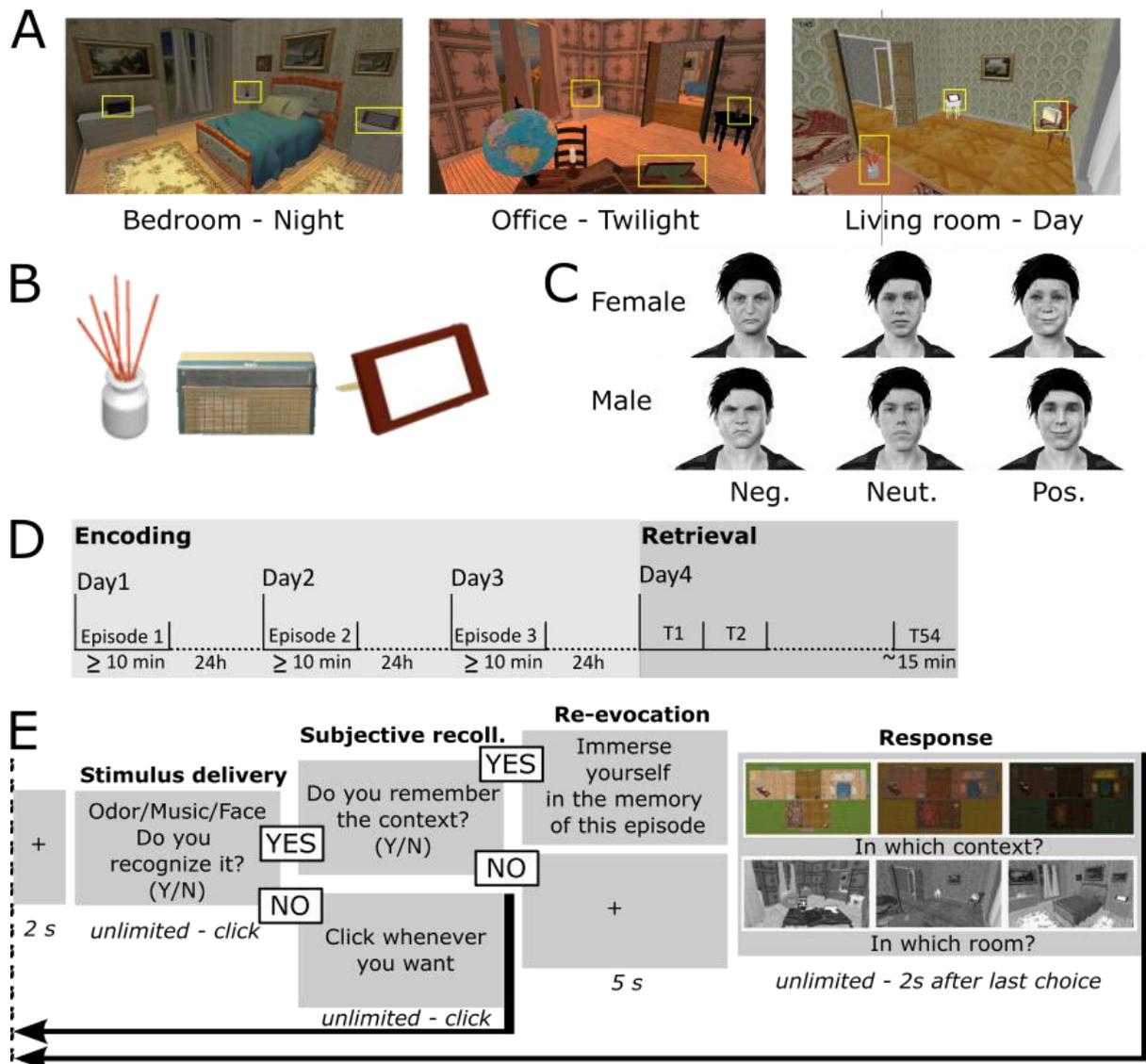


Figure 1 Experimental protocol. (A) The three rooms of the virtual house in three different periods of the day (Bedroom in daylight, Office in twilight, Living-room in nightlight) to illustrate examples of the encoding situations. Yellow rectangles represent the devices with which the stimuli were presented (only added for illustration purposes here). (B) Interactive devices: a fragrance diffuser for odors, a radio for musical excerpts, and a frame for face pictures. (C) Example of face stimuli as a function of gender and a priori emotional category (for an example of music stimuli, see suppl material). (D) The temporal course of the encoding and the retrieval sessions. During the encoding, participants discovered one episode a day over three days. On the fourth day, the memory of the episodes was tested. (E) Example of a trial of the retrieval phase. After a short period of rest, participants were told which stimulus was presented, either an odor, a musique or a face and they must decide whether the cue was a target or a distractor (recognition). In case of a 'No' response, the participants must click on the trackball and the trial ended. In case of a 'Yes' response, the participants were required to indicate their subjective recollection. A 'No' response was followed by 5s of rest. A 'Yes' response was followed by 5s of re-evocation. Then, participants selected a period of time (context) and a room (location) in which this cue was presented at encoding and the trial ended.

2.4.3 Sensory evaluations

At the end of the experiment, participants were required to rate the 54 stimuli (target and distractors) in terms of pleasantness (unpleasant – neutral – pleasant), emotional intensity (very weak – very intense), familiarity (unknown – very familiar), and motivation or wanting (“Do you want to explore again this stimulus?” Not at all – extremely) using non-graduated scales presented with an in-house LabView software. The pleasantness scale was divided into two equal parts by a “neutral” value separating the ratings of unpleasantness and pleasantness. Each unimodal stimulus was presented for about 6 s and participants had been given 5 s to respond to each question. They were trained with a short practice session of one item per sensory modality (*i.e.*, additional items that were not part of the present experiment and were emotionally neutral).

2.5 Data analyses

2.5.1 Episodic memory performance

Episodic memory retrieval has been divided into several steps. It first required participants to accurately recognize target cues (step “Cue+”) from all cues presented (“Cue”) and then to choose the accurate period and the accurate room (accurate period only: “Period+”; accurate room only: “Room+”; both: “Period+ Room+”) in which the cue had been presented during encoding (Figure 2A). Based on the ability of the participants to reach each step or not, memory episodic responses were organized into categories. For recognition responses, four response categories were defined from the experimental conditions (target vs. distractor) and the participants' behavioral responses (“Yes” vs. “No”) based on the signal detection theory (Lockhart and Murdock, 1970): *Hit* and *Miss* occurred when the target items were correctly recognized or incorrectly rejected, respectively, and correct rejection (*CR*) and false alarm (*FA*) occurred when the distractor items were correctly rejected or incorrectly recognized, respectively. For episodic responses, based on the period and room selected, four response categories were defined: *Room-Period-* when both inaccurate room and period were selected, corresponding to a completely incorrect episodic response, *Room- Period+* when the inaccurate room but accurate period were selected and *Room+Period-* when the accurate room but inaccurate period were selected, corresponding both to partially correct episodic response, and *Room+ Period+* when the participants selected both the accurate period and room, corresponding to a complete correct episodic response.

EM performance was computed as the probability to reach each episodic step (Figure 2A): (1) recognize the memory cue among distractors (from Cue to Cue+; 1 response out of 2, “Yes” or “No”; *i.e.*, 50% chance to pick the correct answer), and in case of recognition, to remember (2) either the Room or the Period (from Cue+ to Room+ or from Cue+ to Period+; 1 response

out of 3; *i.e.*, 33% chance to select the correct answer) and then (3) the Period or the Room (from Room+ or Period+ to Room+Period+; 1 response out of 3; *i.e.*, 33% chance to select the correct answer) in which the cue was initially encoded (Figure 2A). EM global performance was evaluated in computing the probability to fully remember the episodes (4) from the cue presentation (from Cue to Room+Period+; $1/2 * 1/3 * 1/3$; *i.e.*, 5% chance), and (5) from the cue recognition (from Cue+ to Room+Period+; $1/3 * 1/3$, *i.e.*, 11% chance). A global performance of stimulus associativity was also evaluated by the probability to retrieve any episodic dimension (*i.e.*, only the room or the period or both) (6) from the cue presentation (from Cue to any episodic retrieval; $1/2 * ((1/3 * 2/3 * 2) + (1/3 * 1/3))$, *i.e.*, 28% chance) or (7) from the cue recognition (from Cue+ to any episodic retrieval ($(1/3 * 2/3 * 2) + (1/3 * 1/3)$), *i.e.*, 56% chance). EM performance was analyzed as a function of 1) Sensory Modality (odor, music, face), 2) Number of Presentation and Encoding Day (day1, day2, day3) (supplementary data), 3) Sensory Evaluation based on participants' individual ratings and sensory modality, and 4) EM performance was also compared to chance levels.

Additionally, recognition performance was resumed with two metrics: The memory score (*d'*) reflects the participant's ability to discriminate between target and distractor items (the higher the value, the better the performance), and response bias score (*c*) reflects participant's tendency to give more "Yes" (positive score) or "No" (negative score) response regardless of experimental condition.

2.5.2 Subjective recollection

Subjective recollection responses were analyzed as the probability to give a "Yes" response (1 "Yes" or "No" response out of 2, *i.e.*, 50 % chance). It was analyzed as a function of Sensory Modality (odor, music, face) and of subsequent episodic memory responses (Room-Period-, Room+Period-, Room-Period+, Room+Period+).

2.5.3 Sensory evaluations

Pleasantness, emotional intensity, familiarity, and wanting ratings of the stimuli were *a posteriori* transformed into scores from 0 to 10. They were analyzed as a function of Sensory Modality (odor, music, face). The relationship between the sensory evaluations of the cues and the episodic memory performance was investigated by analyzing the effect for each sensory evaluation (Pleasantness, Emotional intensity, Familiarity, Wanting) on each of the 7 steps of interest of the episodic memory process.

2.6 Statistical analyses

2.6.1 Linear and generalized linear mixed effect models

2.6.1.1 General method

Most of the analyses were conducted with linear and generalized linear mixed effect models (LMM and GLMM, respectively). A more detailed presentation of these models and explanations about the reasons that motivated this choice was presented in (Rey *et al.*, a). The general analyses approach was as follows: Mixed models were fitted with the “mixed” function from the *afex* library (Bates *et al.*, 2015; Singmann *et al.*, 2021). These mixed models were fitted by the Laplace approximation method. The global procedure used to construct models was the following: unless specified, the maximal model justified by the design and the number of datapoints was constructed first (Barr *et al.*, 2013). In case of singular fit or convergence errors, random effects were selected (reduced) by comparing nested models with likelihood ratio test (LRT) and a backward selection heuristic, so the model complexity was reduced until the statistical test was significant (same method as in Matuschek *et al.*, 2017) or no more errors were noted. As the experiment had been constructed with a within-participant design, the random Participant intercept was never dropped. LMM were checked using package *performance* (Lüdtke *et al.*, 2021) from the *easystat* collection (Makowski *et al.*, 2020). GLMM were checked using the *DHARMA* package (Hartig, 2021) and *p*-values were computed by LRT. *P*-values were computed using the package *lmerTest* (Kuznetsova *et al.*, 2015) by the Kenward-Roger approximation with type 2 tests when the model did not contain interactions, and type 3 tests when it did.

2.6.1.2 Sensory modality effects on episodic memory

The first step of episodic memory, which is defined here as recognition responses, has been analyzed by adapting the signal detection theory framework to GLMM analysis (Fawcett and Ozubko, 2016; Wright and London, 2009; Wright *et al.*, 2009), using sum contrast coding so that the intercept represents reference modalities. A probit mixed model (estimated using maximum likelihood or ML, and base optimizer) was fitted to predict the probability to answering “Yes” with a 3 Sensory Modality (odor, music, face) x 2 Item Type (target, distractor) mixed design. The maximal random effect structure was: (Item Type || participant) + (Item Type || Item). Item Type was centered, applying the following dummy coding: distractors were coded as -0.5 and targets were coded as 0.5. For Sensory Modality, the reference level was “face”, and for Item Type it was 0. This coding allowed us to estimate *d'* and *c* scores for each sensory modality using the estimates calculated by the model. Item Type estimate represented the propensity to discriminate between old and new face, as face was the reference level. It was thus a measure of face's *d'*. For odor and music, this value was equal to the value for face plus

the estimate of the interaction between Sensory Modality and Item Type for odor and music (*i.e.*, for odor, “ d'_{face} ” estimate + “Sensory Modality [odor] * Item Type” estimate). As the intercept represents the propensity to say “Yes” for a face with Item Type and Version at their mean level, its probit-transformed estimates were equal to the bias score (c) for faces. Bias scores for odor and music were then calculated by adding their respective coefficient to the intercept value (*i.e.*, for odor, “ c_{face} ” estimate + “Sensory Modality [odor] * Item Type” estimate). This model allowed us also to compute the probability to reach Cue+ from Cue.

For the other episodic steps of interest (from Cue+ to Room+ and then to Room+Period+, from Cue+ to Period+ and then to Room+Period+, from Cue to any episodic retrieval, from Cue+ to any episodic retrieval), a logistic mixed model was modeled in order to predict the probability to succeed in each level with a 3 Sensory Modality (odor, music, face) mixed design using logit link and estimated with ML and base optimizer. The maximal random effect structure was the same for all models: (Sensory Modality | Participant) + (1 | Item). Random effect structure for all the final models differed for each episodic step.

2.6.1.3 Subjective recollection

The maximal logit mixed model was constructed with Sensory Modality (odor, music, face) and Episodic Score (Room-Period-, Room+Period-, Room-Period+, Room+Period+) and their interaction as fixed effect. The maximal random effect structure was: (Sensory Modality*Episodic_Score| Participant) + (Episodic Score | Item).

2.6.1.4 Sensory evaluations

For each sensory evaluation (pleasantness, emotional intensity, familiarity, and wanting) a linear mixed model was fitted (estimated using restricted maximum likelihood, or REML, and *nloptwrap* optimizer) with a 3 Sensory Modality (odor, music, face) x 3 *a priori* Emotional Category (negative, neutral, positive) mixed design. The maximal random effect structure was (Sensory Modality * Emotional Category | Participant) + (Emotional Category | Item). Random effect structure for all the final models and transformations of dependent variable differed for each sensory evaluation.

2.6.1.5 Effect of sensory evaluations on memory performance

For each of the episodic steps of interest, a model was fitted to predict the probability to succeed, with as fixed effects one sensory evaluation (pleasantness, emotional intensity, familiarity, and wanting) and its interaction with the sensory modality. Pleasantness evaluation was added in its quadratic form. The maximal random effect structure was (Sensory Modality: Sensory Evaluation + Sensory Modality | Participant) + (Sensory Evaluation | item). It was then reduced to minimize convergence errors.

2.6.2 General statistical information

Statistical analyses were performed in R studio with R 4.0.4 (R Development Core Team, 2021), with orthogonal sum-to-zero contrasts set in the default settings. Mixed models and ANOVAs were fitted with the *afex* library version 0.28-1 and based on *lme4* package. When fixed effects were significant, *post-hoc* Bonferroni comparisons were run to determine significant differences among the estimated marginal means (EMM) with *emmeans* package version 1.5.5-1 (Lenth *et al.*, 2020). The same package was used to test the difference with chance value. The significance level was set to $p < .05$, and a p -value between 0.05 and 0.09 was referred to as a tendency. For clarity purpose, only tests addressing the main goal of the present paper were reported. Similarly, when interaction between effects was statistically significant, effects of the main factors were not reported. When interaction between variables were tested, the trend of the slope was reported. Data were described with estimated marginal means and its standard error of the mean (EMM \pm SEM).

3 Results

3.1 Memory performance

Memory performance are reported on Figure 2B and 2C.

3.1.1 From Cue to Cue+

Hit (*i.e.*, accurate recognition of target cues) and FA (*i.e.*, inaccurate recognition of distractor cues) recognition responses were analyzed with the [(Response category | Participant) + (Response category | Item)] random effect structure (conditional $R^2 = 0.62$, marginal $R^2 = 0.62$). The analysis yielded significant main effects of Sensory Modality ($\chi^2(2) = 12.71$, $p < .01$) and Response category ($\chi^2(1) = 55.10$, $p < .001$). A significant interaction between Sensory Modality and Response Category ($\chi^2(2) = 32.30$, $p < .001$) showed that the probability for a Hit was significantly higher for music (EMM \pm SEM, 96.7 ± 1.3 %) than for face (73.1 ± 4.3 %; z 's = 6.45, p 's < .001) and for odor (78.4 ± 3.9 %; z 's = 5.55, p 's < .001), and that the probability for an FA response was higher for odor (15.7 ± 3.3 %) than for music (5.3 ± 1.7 %; $z = 3.36$, $p < .01$). The FA probability associated to face was not different (10.1 ± 2.6 %) from the FA probability for both odor and music. In other words, odor and face cues evoked less accurate recognition, but only odor cues evoked more inaccurate recognition than music and/or face cues. Moreover, the probability to obtain Hit and FA recognition responses differed significantly from random responses (50 %), indicating relatively good recognition performance of cues from all three sensory modalities (z 's ≥ 5.34 , p 's < .001). Memory score d' was estimated at 1.80 ± 0.46 for odors, 3.47 ± 0.49 for music, and 1.90 ± 0.20 for faces. Bias score c was estimated at -0.11 ± 0.20 for odors, -0.01 ± 0.22 for music, and -0.33 ± 0.10 for faces.

3.1.2 From Cue+ to Room+ to Room+Period+

3.1.2.1 From Cue+ to Room+

The model was fitted with maximal random-effect structure (conditional $R^2 = 0.07$, marginal $R^2 = 0.02$). This probability was significantly influenced by the Sensory Modality of the cue ($\chi^2(2) = 6.33$, $p < .05$), such that it was higher for odors ($45.5 \pm 4.3\%$) than for music ($30.9 \pm 3.6\%$; $z = 2.54$, $p < .05$), with no difference between face ($34.7 \pm 4.1\%$) and both odor ($z = 1.79$, $p = .22$) and music ($z = 0.72$, $p = 1.00$) cues. Moreover, this probability was significantly higher than chance (33 %) for odor ($z = 2.82$, $p < .05$), but not for music ($z = 0.69$, $p = 1.00$) and face ($z = 0.33$, $p = 1.00$) cues.

3.1.2.2 From Room+ to Room+Period+

The model was fitted with [(1 | Participant)] random-effect structure (conditional $R^2 = 0.15$, marginal $R^2 = 0.02$). This probability was not significantly affected by the Sensory Modality of the cue ($\chi^2(2) = 3.89$, $p = .14$). The global probability ($41.5 \pm 4.6\%$) tended to be higher than chance (33 %; $z = 1.77$, $p = .078$).

3.1.3 From Cue+ to Period+ to Room+Period+

3.1.3.1 From Cue+ to Period +

The model was fitted with the [(1 | Participant) + (1|stimulus_name)] random effect structure (conditional $R^2 = 0.06$, marginal $R^2 = 0.02$). This probability was significantly influenced by the Sensory Modality of the cue ($\chi^2(2) = 8.8$, $p < .05$), such that it was higher for both odor ($45.2 \pm 3.9\%$) and face ($43.9 \pm 4.0\%$) than for music ($32.1 \pm 3.2\%$; z 's ≥ 2.43 , p 's $< .05$) cues, without any differences between faces and odors ($z = 0.24$, $p = 1.00$). Moreover, this probability was significantly higher than chance (33 %) for odor and face (z 's ≥ 2.63 , p 's $< .05$), but not for music ($z = 0.38$, $p = 1.00$) cues.

3.1.3.2 From Period+ to Room+Period+

The model was fitted with [(1 | Participant) + (1|stimulus_name)] random-effect structure (conditional $R^2 = 0.12$, marginal $R^2 = 0.01$). This probability was not significantly affected by the Sensory Modality of the cue ($\chi^2(2) = 1.39$, $p = .50$). The global probability ($38.7 \pm 4.1\%$) was not significantly different than chance (33 %; $z = 1.29$, $p = .20$).

3.1.4 From Cue to Room+Period+

The model was fitted with [(1 | Participant) + (1 | Item)] random effect structure (conditional $R^2 = 0.12$, marginal $R^2 = 0.01$). This probability was not affected by the Sensory Modality of the

cue ($\chi^2(2) = 2.39, p > .30$). The global probability ($11.4 \pm 1.7\%$) was significantly higher than chance (6% ; $z = 3.44, p < .001$).

3.1.5 From Cue+ to Room+Period+

The model was fitted with [(1 | Participant) + (1 | Item)] random effect structure (conditional $R^2 = 0.12$, marginal $R^2 = 0.02$). This probability was significantly affected by the Sensory Modality of the cue [$\chi^2(2) = 7.16, p < .05$], such that it was higher for odor ($18.9 \pm 3.2\%$) than for music ($10.4 \pm 2.1\%$; $z = 2.71, p < .05$), with no difference between face ($15.8 \pm 3.0\%$) and both odor ($z = 0.84, p = 1.00$) and music ($z = 1.8, p = .21$) cues. Moreover, this probability was significantly higher than chance (11%) for odor ($z = 2.40, p < .05$), but not for music ($z = 0.34, p = 1.00$) and face ($z = 1.55, p = .37$) cues.

3.1.6 From Cue to any episodic retrieval

The model was fitted with [(Sensory Modality | Participant)] random-effect structure (conditional $R^2 = 0.03$, marginal $R^2 = 0.01$). This probability was not significantly affected by the Sensory Modality of the cue ($\chi^2(2) = 4.66, p = .10$). The global probability ($48.8 \pm 2.1\%$) was significantly higher than chance (28% ; $z = 9.94, p < .001$).

3.1.7 From Cue+ to any episodic retrieval

The model was fitted with maximal random-effect structure (conditional $R^2 = 0.05$, marginal $R^2 = 0.04$). This probability was significantly influenced by the Sensory Modality of the cue ($\chi^2(2) = 10.95, p < .01$), such that it was higher for odor ($72.1 \pm 3.8\%$) than for music ($52.6 \pm 3.5\%$; $z = 3.37, p < .01$), with no difference between face ($61.9 \pm 4.1\%$) and both odor ($z = 1.73, p = .25$) and music ($z = 1.86, p = .20$) cues. Moreover, this probability was significantly higher than chance (56%) for odors ($z = 4.39, p < .001$), but not for music ($z = 0.85, p = 1.00$) and face ($z = 1.56, p = .35$) cues.

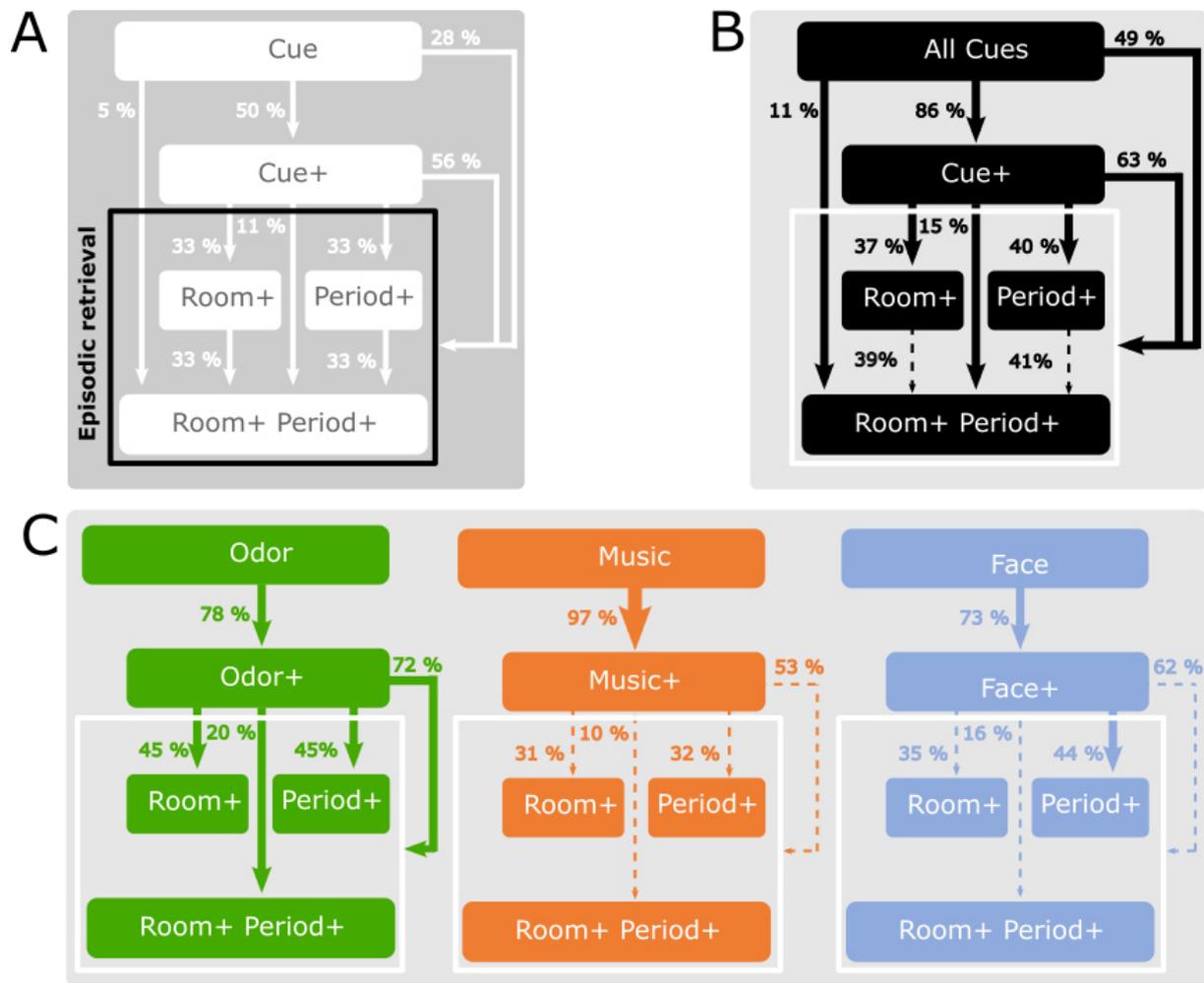


Figure 2. Episodic memory performance. A. Chance level value for each probability to go from one step to another. B. Probability to go from one step to another when no sensory modality effect was significant. C. Sensory-modality specific probability to go from one step to another. Differences of arrow thickness among a same step represent significant difference of probability between sensory modalities. Solid lines represent significant probability; dashed lines represent non-significant probability. When sensory modality effect was not significant no arrow were displayed.

3.2 Subjective recollection

Metacognition was evaluated, as participants were asked whether they felt they could remember the whole episode. The probability to answer “Yes” was represented in Figure 3. It was fitted with a model containing (1 | Participant) random effect structure (conditional $R^2 = 0.18$, marginal $R^2 = 0.08$). The probability to answer “Yes” tended to differ with Sensory Modality ($\chi^2(2) = 5.03$, $p = .081$), however, with no significant differences between sensory modalities (z 's < 2.07, p 's > .12). Moreover, the probability to answer “Yes” was higher than 50 % for odors (63.70 ± 4.51 %; $z = 3.04$, $p < .01$) but not faces (52.24 ± 4.87 %; $z = 0.46$, $p = 1.00$), neither music (54.41 ± 4.50 %; $z = 0.98$, $p = .98$). The probability to answer “Yes” significantly differed for Episodic Score ($\chi^2(2) = 30.47$, $p < .001$), being higher before accurate

episodic responses (Room+Period+; 76.04 ± 4.74 %) than before inaccurate episodic responses (Period-Room-; 44.1 ± 4.3 %; $z = 5.07$, $p < .001$), and than before both partial episodic response (Room+Period-, 56.6 ± 5.2 %; $z = 2.97$, $p < .05$; Room-Period+; 48.2 ± 4.9 %; $z = 4.25$, $p < .001$). Moreover, this probability was above chance level (50 %) when followed by complete episodic response only (Room+Period+, $z = 5.50$, $p < .001$; other responses, z 's < 1.38 , p 's > 0.67).

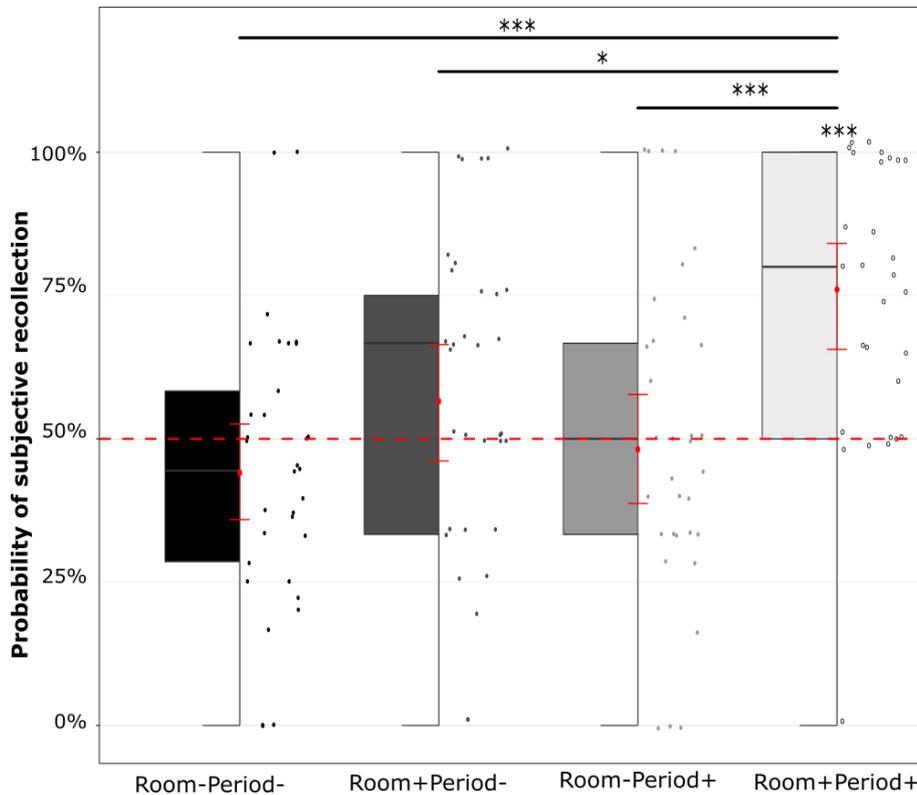


Figure 3. Subjective recollection responses. Probability to give a “Yes” response as a function of subsequent episodic memory response. Individual raw data are represented with dots. The distribution of data is displayed with boxplot (minimum, first quartile, median, third quartile, maximum) in black. The model estimated means and their dispersion (SEM) are represented in red. The dashed horizontal line indicates the random level. * $p < .05$; *** $p < .001$.

3.3 Sensory evaluations and their influence on memory performance

3.3.1 Sensory evaluations

Pleasantness evaluations were analyzed with the [(Emotional Category * Sensory Modality | Participant) + (1 | Item)] random effect structure (conditional $R^2 = 0.59$, marginal $R^2 = 0.51$). The analysis yielded a significant interaction between Pleasantness and Sensory Modality ($F_{(4,49)} = 2.67$, $p < .05$). The interaction indicated that for odor cues (negative: 2.66 ± 0.33 ; neutral: 5.13 ± 0.38 ; pleasant: 6.97 ± 0.29), music cues (negative: 3.97 ± 0.47 ; neutral: 6.72 ± 0.31 ; pleasant: 7.62 ± 0.35) and face cues (negative: 1.26 ± 0.29 ; neutral: 4.69 ± 0.33 ; pleasant: 6.78 ± 0.41), negative cues were judged as less pleasant than neutral and positive

cues. Neutral cues were also judged less pleasant than positive cues (z 's > 4.18 , p 's $< .001$), except for neutral music where the difference from positive music only tended to be significant ($z = 2.32$; $p = .075$). Negative faces were more negative than negative odors ($z = 3.25$; $p < .01$) and music ($z = 5.27$; $p < .001$), and neutral music was more positive than neutral odors ($z = 3.36$; $p < .01$) and faces ($z = 4.05$; $p < .01$).

Emotional intensity evaluations (transformed with a power function (exponent 1.2) to avoid heteroskedasticity and non-normality) were analyzed with the [(Sensory Modality: Emotional Category + Emotional Category | Participant) + (1 | Item)] random effect structure (conditional $R^2 = 0.34$, marginal $R^2 = 0.09$) and the *Bobyqa* optimizer. The analysis yielded a significant main effect of Sensory Modality ($F_{(2,46)} = 10.03$, $p < .001$), indicating that odors (6.26 ± 0.249) were judged as being more strongly emotional than face (5.02 ± 0.27 ; $t_{(46)} = 4.50$, $p < .001$) and tended to be more emotional than music (5.65 ± 0.23 ; $t_{(48)} = 2.42$, $p = .058$) cues. No significant difference of emotional intensity was observed between face and music ($t_{(46)} = 2.22$, $p = .09$) cues. The analysis also yielded a significant main effect of Emotional Category ($F_{(2,46)} = 4.91$, $p < .05$), indicating that positive stimuli (5.91 ± 0.24) were judged as being more strongly emotional than neutral ones (5.29 ± 0.21 ; $t_{(49)} = 2.83$, $p < .05$). No significant difference of emotional intensity was observed between positive and negative (5.74 ± 0.29) cues ($t_{(42)} = 0.54$, $p = 1.00$) and between neutral and negative cues ($t_{(47)} = 1.70$, $p = .29$).

Familiarity evaluations were analyzed with the [(Sensory Modality: Emotional Category + Sensory Modality | Participant)] random effect structure (marginal $R^2 = 0.12$) and the *bobyqa* optimizer. The analysis yielded a significant main effect of Sensory Modality ($F_{(2,29)} = 13.62$, $p < .001$), indicating that odors (6.03 ± 0.26) were judged as being more familiar than music (5.06 ± 0.20) and face (4.75 ± 0.33 ; t 's ≥ 4.44 , p 's $< .001$) cues. No significant difference of familiarity was observed between faces and music ($t_{(30)} = 1.12$, $p = .82$). The analysis also yielded a significant main effect of Emotional Category ($F_{(2,29)} = 14.22$, $p < .001$), indicating that negative stimuli (4.39 ± 0.31) were judged as being less familiar than neutral (5.61 ± 0.23) and positive (5.84 ± 0.21 ; t 's ≥ 4.60 , p 's $< .001$) stimuli. No significant difference of familiarity was observed between neutral and positive ($t_{(30)} = 1.15$, $p = .78$) stimuli.

Wanting evaluations were analyzed with the [(Sensory Modality: Emotional Category + Sensory Modality | Participant) + (1 | Item)] random effect structure (marginal $R^2 = 0.40$) and the *Bobyqa* optimizer. The analysis yielded a significant main effect of Sensory Modality ($F_{(2,46)} = 11.42$, $p < .001$), indicating that music (5.28 ± 0.26) was judged as being more wanted than odor (4.40 ± 0.24 ; $t_{(46)} = 2.87$, $p < .05$) and face (3.73 ± 0.20 ; $t_{(46)} = 4.82$, $p < .001$) cues. Odors tended to be more wanted than faces ($t_{(47)} = 2.36$, $p = .067$). The analysis also yielded a significant main effect of Emotional Category ($F_{(2,44)} = 54.16$, $p < .001$), indicating that negative

(2.36 ± 0.24), neutral (4.77 ± 0.20) and positive (6.28 ± 0.27) stimuli were judged to be different between each other's (t 's ≥ 5.66 , p 's $< .001$): negative stimuli were less wanted than neutral stimuli that were also less wanted than positive stimuli across sensory modalities.

3.3.2 Influence of sensory evaluations on memory performance

The following analyses tested the effect on individual sensory evaluations of the cues (pleasantness, emotional intensity, familiarity, wanting) on each of the steps of the episodic memory process whose probability was significantly higher than chance. Results are presented on Figure 4.

3.3.2.1 From Cue to Cue +

Pleasantness. The model was fitted with [(1 | Participant) + (1 | Item)] as random effects (conditional $R^2 = 0.31$, marginal $R^2 = 0.03$) and the *Bobyqa* optimizer. A significant principal effect of quadratic pleasantness ($\chi^2_{(2)} = 7.70$, $p < .05$) revealed that pleasantness influenced the target recognition score in a U-shaped fashion, with both negative and positives cues favoring recognition, in contrast with neutral cues.

Emotional Intensity. The model was fitted with [(1 | Participant) + (1 | Item)] as random effects (conditional $R^2 = 0.34$, marginal $R^2 = 0.22$) and the *nlminbwrap* optimizer. A significant interaction between Emotional Intensity and Sensory Modality ($\chi^2_{(2)} = 35.55$, $p < .001$) revealed that the slope for music cues (0.38 ± 0.07) was significantly different than the slope for odors (0.08 ± 0.05 , $z = 4.81$, $p < .001$) and faces (0.01 ± 0.05 , $z = 5.84$, $p < .001$). No significant difference of emotional intensity slope was observed between odors and faces ($z = 1.63$, $p = .31$). Moreover, the emotional intensity slope for music was significantly different from zero ($z = 5.25$, $p < .001$) and tended to be different from zero for odors ($z = 1.71$, $p = .087$), but not for faces ($z = 0.25$, $p = .80$).

Familiarity. The model was fitted with [(1 | Participant) + (1 | Item)] as random effects (conditional $R^2 = 0.33$, marginal $R^2 = 0.20$) and the *Bobyqa* optimizer. A significant interaction effect between Familiarity and Sensory Modality ($\chi^2_{(2)} = 26.52$, $p < .001$) revealed that the slope for music (0.37 ± 0.07) was significantly different from the slope for odors (0.07 ± 0.04 , $z = 4.71$, $p < .001$) and faces (0.05 ± 0.04 , $z = 4.96$, $p < .001$) cues. No significant difference of emotional intensity slope was observed between odors and faces ($z = 0.45$, $p = 1.00$). Moreover, the emotional intensity slope for music was significantly different from zero ($z = 5.37$, $p < .001$) and tended to be different from zero for odors ($z = 1.71$, $p = .087$), but not for faces ($z = 1.24$, $p = .22$).

Wanting. The model was fitted with [(1 | Participant) + (1 | Item)] as random effects (conditional $R^2 = 0.32$, marginal $R^2 = 0.15$). A significant interaction between Wanting and Sensory modality

($\chi^2(2) = 20.79, p < .001$) revealed that the slope for music (0.26 ± 0.06) was significantly different from the slope for odor (0.00 ± 0.04) and face ($-0.03 \pm 0.05, z$'s $>3.99, p$'s $< .001$) cues. No significant difference of wanting slope was observed between odors and faces ($z = 0.59, p = 1.00$). Moreover, the wanting slope for music was significantly different from zero ($z = 4.16, p < .001$) but not for odor and faces (z 's $< 0.63, p$'s $> .52$).

3.3.2.2 From Cue+ to Room+

Pleasantness. The model was fitted with [(Pleasantness || Participant) + (Pleasantness || Item)] as random effects (conditional $R^2 = 0.01$, marginal $R^2 = 0.01$). It revealed no significant effects.

Emotional Intensity. The model was fitted with [(Sensory Modality : Emotional Intensity | Participant) + (0 + Emotional Intensity || Item)] as random effects (conditional $R^2 = 0.02$, marginal $R^2 = 0.02$) and the *Nelder-Mead* optimizer. It revealed no significant effects.

Familiarity. The model was fitted with [(Sensory Modality : Familiarity | Participant) + (Familiarity | Item)] as random effects (conditional $R^2 = 0.15$, marginal $R^2 = 0.02$) and the *Nelder-Mead* optimizer. A significant interaction effect between Familiarity and Sensory Modality ($\chi^2(2) = 8.02, p < .05$) revealed that the slope for odors (0.06 ± 0.04) was significantly different from the slope for music ($-0.03 \pm 0.04, z = 2.71, p < .05$), with no other differences (z 's $< 2.17, p$'s $> .09$). The familiarity slopes for all stimuli were not different from zero (z 's $< 1.67, p$'s > 0.09).

Wanting. The model was fitted with [(1 | Participant) + (1 || Item)] as random effects (conditional $R^2 = 0.10$, marginal $R^2 = 0.02$). A significant interaction effect between Wanting and Sensory Modality ($\chi^2(2) = 6.06, p < .05$) revealed that the slope for odors (0.10 ± 0.04) tended to be significantly different from the slope for music ($-0.01 \pm 0.04, z = 2.20, p = .084$), with no other differences (z 's $< 2.08, p$'s $> .11$). The slope was different from 0 for odors only ($z = 2.51, p < .05$; for music and faces, z 's = $2.08, p$'s $> .11$)

3.3.2.3 From Cue+ to Period+

Pleasantness. The model was fitted with [(1 | Participant) + (1 | Item)] as random effects (conditional $R^2 = 0.05$, marginal $R^2 = 0.01$). It revealed no significant effects.

Emotional Intensity. The model was fitted with [(1 | Participant) + (1 | Item)] as random effects (conditional $R^2 = 0.05$, marginal $R^2 = 0.02$). A significant interaction effect between Emotional intensity and Sensory modality ($\chi^2(2) = 7.89, p < .05$) revealed that the slope for faces (0.08 ± 0.05) was significantly different from the slope for music ($-0.02 \pm 0.04, z = 2.72, p < .05$), with no other differences (z 's $< 1.98, p$'s $> .14$). Moreover, the slope for faces was the only slope

that tended to be different from zero ($z = 1.73$; $p = .084$; z 's < 1.19 , p 's > 0.23 for odor and music).

Familiarity. The model was fitted with [(1 | Participant) + (1 | Item)] as random effects (conditional $R^2 = 0.06$, marginal $R^2 = 0.02$). It revealed no significant effects.

Wanting

The model was fitted with [(1 | Participant) + (1 | Item)] as random effects (conditional $R^2 = 0.06$, marginal $R^2 = 0.02$). A significant interaction effect between Wanting and Sensory modality ($\chi^2(2) = 7.61$, $p < .01$) revealed that the slope for odor (0.07 ± 0.03) was significantly different from the slope for music (-0.01 ± 0.03 , $z = 2.45$; $p < .05$) cues with no other significant differences (z 's < 2.17 , p 's > 0.09). Moreover, the slope for odors was significantly differed from zero ($z = 2.15$; $p < .05$), the slope for faces (0.07 ± 0.04) tending to be different from zero ($z = 1.85$, $p = .064$), but not the slope for music ($z = 0.36$, $p = .72$)

3.3.2.4 From Cue to Room+Period+

Pleasantness. The model fitted with [(1 | Participant) + (1 | Item)] as random effects (conditional $R^2 = 0.17$, marginal $R^2 = 0.03$) and the *Bobyqa* optimizer. A significant principal effect of quadratic Pleasantness ($\chi^2(2) = 8.35$, $p < .05$) revealed that pleasantness of the cues improved in a u-shaped fashion the probability to retrieve accurately the whole episode after perceiving the cue.

Emotional Intensity. The model was fitted with [(1 | Participant)] as random effects (conditional $R^2 = 0.12$, marginal $R^2 = 0.02$). A significant principal effect of Emotional intensity (0.11 ± 0.05 , $\chi^2(1) = 3.90$, $p < .05$) revealed that emotional intensity of the cues significantly improved the probability to retrieve accurately the whole episode after perceiving the cue.

Familiarity. The model was fitted with [(1 | Participant) + (1 | Item)] as random effects (conditional $R^2 = 0.16$, marginal $R^2 = 0.02$). It revealed no significant effects.

Wanting. The model was fitted with [(Wanting || Participant) + (1 | Item)] as random effects (conditional $R^2 = 0.05$, marginal $R^2 = 0.02$). It revealed no significant effects.

3.3.2.5 From Cue+ to Room+Period+

Pleasantness. The model was fitted with [(1 | Participant) + (1 | Item)] as random effects (conditional $R^2 = 0.15$, marginal $R^2 = 0.03$) and the *Bobyqa* optimizer. A significant principal effect of quadratic Pleasantness ($\chi^2(2) = 6.34$, $p < .05$) revealed that pleasantness improved in a U-shaped fashion the probability to retrieve accurately the whole episode after accurately recognizing the cue.

Emotional Intensity. The model was fitted with [(1 | Participant)] as random effects (conditional $R^2 = 0.10$, marginal $R^2 = 0.03$). It revealed no significant effects.

Familiarity. The model was fitted with [(1 | Participant) + (1 | Item)] as random effects (conditional $R^2 = 0.13$, marginal $R^2 = 0.03$) and the *Bobyqa* optimizer. A tendency effect of the interaction between Sensory Modality and Familiarity ($\chi^2_{(1)} = 5.80$, $p = .055$) revealed that stimulus familiarity tended to improve the probability to retrieve accurately at least one dimension. Odor slope (0.08 ± 0.05) tended to be significantly different than music slope (-0.03 ± 0.06); $z = 2.37$, $p = .053$). However, no differences with face slope (0.04 ± 0.05) were observed (z 's < 1.52 , p 's $> .38$). None of these slopes were different from zero (z 's < 1.62 , p 's $> .10$).

Wanting. The model was fitted with [(Wanting || Participant) + (1 | Item)] as random effects (conditional $R = 0.04$, marginal $R^2 = 0.03$) and the *bobyqa* optimizer. A significant interaction between Sensory modality and Wanting ($\chi^2_{(2)} = 6.75$, $p < .05$) revealed that odor slope (0.11 ± 0.05) was significantly different from the music slope (-0.01 ± 0.05 , $z = -2.60$, $p < .05$). Other comparisons were non-significant (face slope, 0.06 ± 0.05 ; z 's < 1.49 , p 's $> .41$). Moreover, the slope was significantly different from 0 for odor ($z = 2.31$, $p < .05$) but not music nor face (z 's < 1.22 , p 's > 0.22) cues.

3.3.2.6 From Cue to any episodic retrieval

Pleasantness. The model was fitted with [(1 | Participant) + (0 +Pleasantness || Item)] as random effects (conditional $R^2 = 0.03$, marginal $R^2 = 0.01$). It revealed no significant effects.

Emotional Intensity. The model was fitted with [(1 | Participant) + (1 | Item)] as random effects (conditional $R^2 = 0.02$, marginal $R^2 = 0.01$) and the *Bobyqa* optimizer. A tendency effect of Emotional Intensity ($\chi^2_{(1)} = 3.68$, $p = .055$) revealed that stimulus emotional intensity tends to improve the probability to retrieve accurately at least one episodic dimension after perceiving the cue (0.061 ± 0.03).

Familiarity. The model was fitted with [(1 | Participant) + (1 | Item)] as random effects (conditional $R^2 = 0.02$, marginal $R^2 = 0.01$). A significant effect of Emotional intensity ($\chi^2_{(1)} = 4.33$, $p < .05$) revealed that stimulus familiarity improved the probability to retrieve accurately at least one dimension after perceiving the cue (0.056 ± 0.03).

Wanting. The model was fitted with [(1 | Participant) + (1 | Item)] as random effects (conditional $R^2 = 0.03$, marginal $R^2 = 0.01$). It revealed no significant effects.

3.3.2.7 From Cue+ to any episodic retrieval

Pleasantness. The model was fitted with [(Pleasantness: Sensory Modality + Pleasantness || Participant) + (Pleasantness || Item)] as random effects (conditional $R^2 = 0.01$, marginal $R^2 = 0.01$). It revealed no significant effects.

Emotional Intensity. The model was fitted with [(Emotional Intensity: Sensory Modality | Participant) as random effects (conditional $R^2 = 0.13$, marginal $R^2 = 0.03$) and the *Nelder-Mead* optimizer. A significant interaction effect between Emotional intensity and Sensory Modality ($\chi^2_{(2)} = 11.84$, $p < .01$) revealed that the slope for odors (0.10 ± 0.05) was significantly different from the slope for music (0.01 ± 0.05 , $z = 2.79$, $p < .05$), with no other differences (z 's < 1.98 ; p 's $> .14$). Moreover, the slope for odors was the only slope that tended to be different from zero ($z = 1.94$, $p = .053$; z 's < 1.21 , p 's > 0.23 for face and music).

Familiarity. The model was fitted with [(1 | Participant) + (1 | Item)] as random effects (conditional $R^2 = 0.03$, marginal $R^2 = 0.02$). A significant interaction effect between Familiarity and Sensory modality ($\chi^2_{(2)} = 8.12$, $p < .05$) revealed that the slope for odors (0.07 ± 0.03) was significantly different than the slope for music (-0.02 ± 0.04 , $z = 3.04$; $p < .01$), with no other differences (z 's < 1.50 , p 's > 0.39). Moreover, the familiarity slope for odors was the only slope that was significantly different from zero ($z = 2.07$; $p < .05$; z 's < 0.65 , p 's > 0.52 for face and music).

Wanting. The model was fitted with [(Wanting || Participant) + (1 | Item)] as random effects (conditional $R^2 = 0.03$, marginal $R^2 = 0.03$). A significant interaction between Wanting and Sensory modality ($\chi^2_{(2)} = 10.12$, $p < .01$) revealed that the slope for odors (0.09 ± 0.04) was significantly different from the slope for music (-0.03 ± 0.03 , $z = 3.26$, $p < .01$), with no other differences (z 's < 1.73 , p 's > 0.25 ; slope for faces, 0.02 ± 0.04). Moreover, the slope for odors was the only slope that was significantly different from zero ($z = 2.47$; $p < .05$; z 's < 0.85 , p 's $> .40$ for face and music).

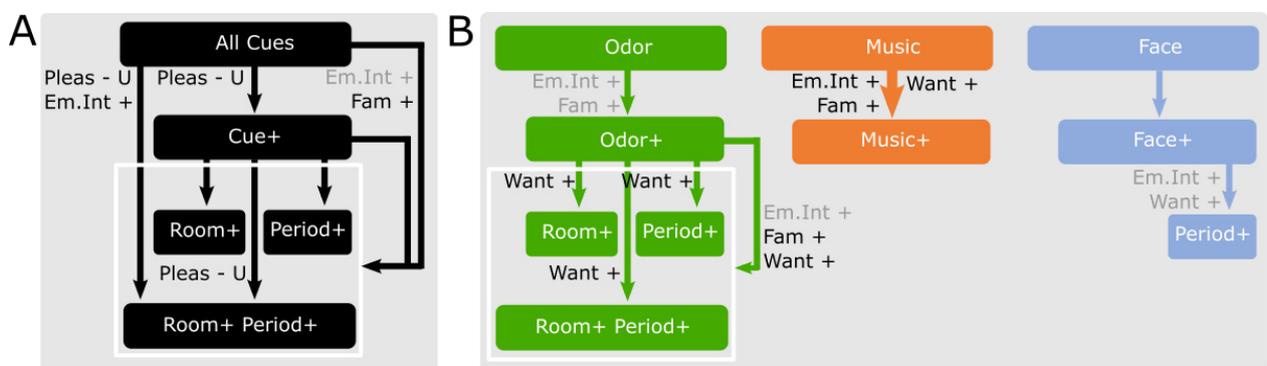


Figure 4. A. Influence of sensory evaluations of the cues on episodic memory performance. Sensory-modality non-specific, and B. Sensory-modality specific positive (+) and quadratic (U) influence of sensory evaluations on the probability to go from one step to another. Em.Int, emotional intensity; Fam, familiarity; Pleas, Pleasantness; Want, wanting. When sensory evaluation labels are in black, $p < .05$, and when in grey, $p < .09$.

4 Discussion

This study was conducted to test the hypothesis that the singularity of odor memory comes from its close link with emotional processes (Hackländer *et al.*, 2018; Larsson *et al.*, 2014). To this aim, episodic memory evoked by odors were compared to episodic memory evoked by music and faces, which have also been reported to be good memory cues (Belfi *et al.*, 2016; Brédart and Barsics, 2012; Jakubowski *et al.*, 2021b) and to have strong links with emotions (Ferreri and Rodriguez-Fornells, 2017; Fuentes-Sánchez *et al.*, 2021; Liu *et al.*, 2014). These emotional stimuli were incidentally discovered by the participants in a protocol allowing for a three-day exploration of episodes displayed in a non-immersive first-person virtual reality. The episodes differed by the period of the day, and the location and the emotional context of a room in a house. The fourth day, participants had to recognize the encoded cues among distractors and to report in which room and period of the day they were discovered. This protocol allowed for the computation of recognition and episodic probabilities and to test the effect of subjective emotional evaluations of the cues on these probabilities. The specificity of our approach was also to consider participants' individual evaluations rather than a priori categorization of the experimental material.

4.1 Recognition performance and episodic memories

Participants were able to correctly recognize the cues among distractors and to retrieve the associated episode, corresponding to the room and the period of the day at encoding. While the recognition memory performance was rather high (86%, chance level 50 %), the episodic memory performance was moderate (11 %, chance level 5 %). Interestingly, participants seemed to have accurate feelings about their memory performance. The complete accurate episodic response (Room+Period+) was the only response for which participants reported more subjective feeling of remembering than an absence of feeling of remembering, and it was the response where the feeling of remembering was the most frequent, in comparison with other episodic responses. These results related to strong memory performance and confidence therein were particularly interesting because the episodes were explored implicitly, and because they were difficult to disentangle, given that the same house was presented during the three days. This difficulty may explain that the results obtained here were poorer than those obtained in the previous study (Rey *et al.*, a), in which participants recollected the whole episode with higher proficiency (24 %). Indeed, in our present study the protocol was different, as participants had to retrieve in which room (defining the context) and where in that room (defining the location) the presented stimulus was presented at encoding. An additional explanation might be given for the slightly decreased performance level here. In the current study, the participants answered by selecting responses from multiple choices, while, in the

previous experiment, they had to go back to the initial encoding context to give their response, which might have allowed the reinforcement of the memory trace each time they responded. Far from being a drawback of the paradigm, the difficulty to remember the whole episode in the present study was coherent with what occurs in real life. Indeed, in real life, the occurrence of a full What-Where-Which memory is rare, and most of the time, all exact elements of the events are not remembered, but only part of them and mostly with memory distortions (Schacter, 1999). Another significant difference with the findings of our previous study was that the probability to retrieve the whole episode after having correctly chosen the period or the room dropped to chance level. This effect might be due to the complexity of the memory that was investigated in our paradigm. The memory tested here was not serial, as it was in our previous study, which unraveled the presence of three underlying memory processes: the recognition of the cue, the memory of the association between either the period or the room, and the episodic memory that gathered the cue, the room and the period. This serial process (investigated in the previous study) may thus be a task of complex source memory, with two levels of location, whereas the process investigated here, more complex, may be closer to a complete episodic memory process.

4.2 Emotional cues benefit episodic memory

To assess the effect of subjective emotion of the cues on the episodic memory performance, we selected *a priori* negative, neutral and positive stimuli, that participants evaluated on 3 properties relative to emotion: pleasantness, emotional intensity and wanting, as well as familiarity. Overall, these evaluations confirmed our *a priori* selection (based on the pretest). Negative, neutral and positive stimuli were rated respectively from less to more pleasant and from less to more wanted. However, neutral stimuli were evaluated as being less emotionally intense than positive stimuli, but not than negative stimuli. Negative stimuli were rated as being less familiar than neutral and positive one, an unsurprising result in light of the well-known link between emotion and familiarity (Delplanque *et al.*, 2008; Distel *et al.*, 1999; Plailly *et al.*, 2011).

Even though the stimuli were also selected to be similar to each other (based on the pretest), we observed differences in sensory evaluation between the stimuli of the different sensory modalities: odors were rated as being more emotionally intense and familiar than music and face stimuli; music stimuli were rated as being more pleasant than the other stimuli and induced more motivation to be re-sampled than odors and faces; and faces tended to be less wanted, less emotionally intense than odor and music cues. While the results of the sensory evaluations concerning odor and music stimuli agreed with our previous findings, faces' sensory evaluations differed with the previous experiment (Rey *et al.*, a), which also used different types of faces. In the current study, faces have been virtualized, and this transformation

influenced their evaluation. This effect may be due to the “Uncanny Valley effect”, explaining that objects imperfectly resembling to a human provoke uncanny feelings of eeriness (MacDorman *et al.*, 2009; Tinwell *et al.*, 2011). The Uncanny Valley effect explains the feeling of strangeness felt when an object, a robot or a character resemble to a real human while keeping non-human attributes. Altogether, these results confirmed that our stimuli selection has been done correctly. However, given the general high inter-individual variability in olfactory perception (Mantel *et al.*, 2019), and given that statistical analyses allow to take this variability into account, we analyzed the data based on the participants judgments and not our a priori categorization aiming to be in the best position to uncover stimuli evaluation effects on memory.

The emotion of the cue had a global effect on almost all memory steps in the current study. Pleasantness was shown to play a major role: positively and negatively valenced emotional cues improving memory performance in comparison to neutral cues. The effect of cue pleasantness also concerned all memory phases, from the perception of a cue and its recognition to its association to the whole episode. This effect on recognition and whole memory was in agreement with our previous findings (Saive *et al.*, 2014a), and with research investigating recognition (Keightley *et al.*, 2011; Larsson *et al.*, 2009; Meng *et al.*, 2017; Stalinski and Schellenberg, 2013) and source memory (Doerksen and Shimamura, 2001; Kensinger and Corkin, 2003; but see: Cook *et al.*, 2017) demonstrating the beneficial influence of valence on memory performance. Emotional intensity of the cues, seemed to favor episodic memory process, with stronger emotional response from the cues evoking more episodic retrieval, but to a lesser extent than the valence dimension. The limited effect of emotional intensity in our experiment may be due to the lack of more emotionally arousing stimuli.

4.3 The influence of the cue’s emotion on recognition memory differed between sensory modalities

Our present findings confirmed that the sensory modality of the cue affects the recognition memory retrieval pattern. We also demonstrated that emotion, evaluated from several sensory dimensions, played an important role in recognition memory retrieval, at both common and different steps for odor, music and face cues.

First, sensory modality influences recognition memory, with music being more easily recognized than both odor and face stimuli. Beyond this difference in absolute recognition performance levels, stimuli from all sensory modalities were well recognized. The odor recognition score (78 %) confirmed previous studies focusing on odor recognition (Larsson *et al.*, 2009; Levy, 2004; Plailly *et al.*, 2019; Saive *et al.*, 2013, 2014a, 2015). The fact that recognition was embedded in a more complex memory task did not seem to affect recognition

memory performance, suggesting its consistency. This lower score for odor recognition than for music recognition may be linked to the difficulty to verbally categorize and communicate about odors (Jraissati and Deroy, 2021; Majid and Burenhult, 2014) and the poor link between odor processed and semantic processes (Olofsson and Gottfried, 2015) that may affect the level with which odors had been processed during encoding (Craik and Lockhart, 1972). Musical clips were recognized with an almost perfect score (97 %), an unsurprising result given the few numbers of target melodies to recognize and the repeated encoding that forced participant to hear the music at least five times. Music may be recognized other basis than odor and face cues, relying on the musical lexicon (Groussard *et al.*, 2019), while odors and virtual faces recognition relies on a more perceptual processes. Face recognition reached a similar performance level than did odor recognition (73 %), as expected with the virtualization of faces that all are presented with the same haircut and outfit (Balas and Pacella, 2015; Crookes *et al.*, 2015; Kätsyri, 2018). These manipulations lead faces to be recognized at a similar performance level as odor, with similar Hit and FA scores.

In addition to showing an influence of the cue's sensory modality, the present study demonstrated that the effect of the cues' sensory modality on memory was dependent on the emotional content of the cue. This suggested that different processes may be involved for odor, music and face recognition. Music was the stimulus for which the recognition was the most strongly influenced by emotion, with the beneficial effect of the cue's emotional intensity and wanting dimensions adding to the valence effect on memory. Notably, our results confirmed previous research on this topic and underlined the importance of taking into account the reward value of musical pieces in memory (Ferreri and Rodriguez-Fornells, 2017; Ferreri *et al.*, 2021). Interestingly, musical excerpts were also the only stimuli for which the recognition was favored by increased feeling of familiarity, which further support the hypothesis suggesting that music recognition is achieved by relying on a musical lexicon (stored in long-term memory), that may involve the sense of familiarity (Groussard *et al.*, 2019).

Odor recognition was also influenced by the cue's emotional intensity and familiarity, but to a lesser extent in comparison with music recognition, as the results failed to reach significance. The fact that this effect was weak was surprising. Odor recognition memory had been shown to be influenced by emotional intensity (Saive *et al.*, 2014a) and by semantic knowledge about odors in general, whether this involved odor familiarity or odor identifiability (Cornell Karnekull *et al.*, 2015; Larsson *et al.*, 2006; Saive *et al.*, 2015). Also, odors were evaluated here as being more emotionally intense and familiar than the musical clips. Taken together this may suggests that odors need to be more emotionally intense and familiar than music in order to improve their recognition. So, these results may reveal that odor recognition might be based to a lesser extent to a familiarity-based recognition and more to a recollection-based recognition.

Face recognition was not affected by emotion and familiarity evaluations. As faces were judged to be the least emotionally intense and wanted stimuli, their low scores may have not allowed for observing an emotional effect. However, it is also probable that faces, due to their virtualization, were too simple to be deeply encoded to be recognized the same way as they were in our previous study (Rey *et al.*, a), where face recognition has been shown to depend on face complexity. As face complexity may promote deeper processing (Craik and Tulving, 1975), the reduced recognition ability for virtualized faces toward real faces is consistent with other research (Balas and Pacella, 2015; Crookes *et al.*, 2015; Kätsyri, 2018). As virtual faces may have been perceived as an other-ethnicity group, it is also possible that this effect strongly reduced face recognition ability (Meissner and Brigham, 2001).

4.4 The influence of the cue's emotion on episodic memory differed between sensory modalities

The present study demonstrated that the sensory modality of the cues influenced episodic memory retrieval and that this influence was conveyed by emotional properties of the cues. The sensory modality effect on episodic memory was more striking than on recognition and appeared after the participant recognized correctly the memory cue. From that moment, participants had to select the period of the day and the room in which the cue had been encountered, which allowed for the computation of four probabilities: the probability to retrieve the whole episode, the probability to retrieve the period, the probability to retrieve the room, and the probability to retrieve at least one dimension (period or room or both). Only odors allowed for the successful achievement of all memory steps, whereas face cues were only able to evoke the associated period, and music cues were only recognized.

Together with our previous study (Plailly *et al.*, 2019; Rey *et al.*, a; Saive *et al.*, 2013, 2014a, 2015) and with other research (Davis, 1977; Herz, 1998; Herz and Cupchik, 1995), these results reflected the strong associative property of odor cues. Odor cues are less recognized and thus less able to evoke numerous memories. However, once recognized their emotional power make them more able to evoke an event or an associated episodic dimension directly. The associative property of odor cues may have been mediated by the cue's emotion. Results indeed underlined the influence of emotion on odor-evoked memory. In particular, emotion influenced all memory steps cued by odors, but not for music or face cues. From the three sensory evaluations reflecting the emotional content of the cues, the motivation to explore again the cue, referred to wanting here, was the only property that affected odor-evoked full memory, while emotional intensity and familiarity affected only the probability to retrieve at least one episodic dimension. The fact that only the wanting dimension was shown to affect odor-evoked memory may be linked to the importance of the attraction/repulsion behavior that

an odor automatically triggers when perceived (Mandairon *et al.*, 2009) and may thus be a the most appropriate measure of odor-induced emotion. There are indeed different measures of a stimulus emotion and they are shown to differently involve brains regions (Jiang *et al.*, 2015). The direct retrieval associated with the odor's emotion may be the factor that make odor memory distinguishable from other ones, a special power of odors.

While being recognized with a high accuracy and being more emotionally salient than faces, and evaluated as more wanted than odor stimuli, music was unable to trigger the retrieval of any dimension of the encoded episode when the cue was recognized,. This result may sustain the hypothesis that musical excerpts were recognized based on a familiarity process (Groussard *et al.*, 2019), and were insufficiently strong to cue other episodic memory dimension. It is possible that the powerful effect of music on autobiographical memory recall (Janata *et al.*, 2007) would be due to lyrics that are often part of a music piece (Halpern *et al.*, 2018). Another explanation may be the lack of ecological validity of the musical excerpts in our paradigm, that were insufficiently long and were too simple to induce reliable emotion and rewarding value that would be strong enough to be associated to an event. Comparing these results with those previously obtained in autobiographical memory experiments (Belfi *et al.*, 2016; Blais-Rochette and Miranda, 2016; Cady *et al.*, 2008; Halpern *et al.*, 2018; Jakubowski *et al.*, 2021b; Janata *et al.*, 2007), it is possible to hypothesize that music might evoke numerous memories by its association to lyrics and its high probability to be recognized. And, the more a cue can be recognized and the more the probability to evoke a memory is high. It is also possible that the recognition of a music cue that is not directly associated to a memory could evoke other memory elements through a generative process.

Face-evoked episodic memory performance was also weak, as faces were able to only induce the retrieval of the associated period. This lack of additional effect may be also due to the weaker link with which faces were associated with emotion in our experiment. Faces were indeed evaluated as being less emotionally intense and wanted than odor and music cues, and the ability of a face to cue the period was found to be enhanced by emotional intensity and wanting. However, due to the difficulty of the task, it may be possible that these emotions, as for music, were insufficiently strong to help faces to evoke the associated episode. As for their recognition ability, it is also possible that the here used, virtual face cues were not sufficiently complex to be associated to episodic dimensions, as a deeper processing promotes the recollection of episodic memory, and as the complexity may promotes a deeper processing (Craik and Tulving, 1975).

4.5 Limitations

The study was designed to ensure a high control on encoding and retrieval conditions, allowing for the computation of precise probabilities and observation of different memory processes. It relied on the what-where-which context definition of episodic memory (Clayton and Dickinson, 1998; Easton and Eacott, 2008b; Tulving, 1972) to investigate the effect of emotion and sensory modality of the cue on recognition and episodic retrieval. However, we did not ask directly participants question about their level of consciousness at retrieval, as Endel Tulving emphasized in his research (Tulving, 1985, 2002). We rather investigated the subjective sensation of having remembered an episode, in a similar but nonequivalent way as it is generally done in remember/known paradigms. The task may thus be viewed as measuring complex source memory, or an episodic-like memory, as it is done in non-human animal studies (Cheke and Clayton, 2013; Clayton and Dickinson, 1998). It might be argued that a potential drawback of the here used paradigm is the lack of ecological validity in the cue choice. It can be indeed interesting to see if the results we obtained become similar for more ecologically valid cues, or cues that are namable. Our stimuli, while selected to be emotional, may suffer from a lack of potential power to be arousing.

4.6 Conclusion

The aim of this study was to test the hypothesis that the cue's evoked emotion influences more strongly episodic memory evoked by odor than episodic memory evoked by music and faces. Our results demonstrated that the whole process underlying the recall of episodic memory was influenced by both the sensory modality and by the emotion of the presented memory cues, the two dimensions interacting with each other. Odors were shown to be powerful memory cues, and the memory they evoked was influenced by their emotional content, in particular their motivational content. Musical clips were specific in reaching high levels of recognition memory, which was influenced by the emotion and familiarity of the music. Faces were recalled at the same level as odor cues, but were only able to evoke the retrieval of the episode period. Thus, this study highlights the power of odors to evoke complex associative memories such as episodic memory, probably thanks to its links with motivational processes.

5 References

- Alonso, I., Dellacherie, D., and Samson, S. (2015). Emotional memory for musical excerpts in young and older adults. *Front Aging Neurosci* 7, 23.
- Annett, J.M. (1996). Olfactory Memory: A Case Study in Cognitive Psychology. *The Journal of Psychology* 130, 309–319.
- Balas, B., and Pacella, J. (2015). Artificial faces are harder to remember. *Computers in Human Behavior* 52, 331–337.
- Barr, D.J., Levy, R., Scheepers, C., and Tily, H.J. (2013). Random effects structure for confirmatory hypothesis testing: Keep it maximal. *Journal of Memory and Language* 68, 255–278.
- Bates, D., Mächler, M., Bolker, B., and Walker, S. (2015). Fitting Linear Mixed-Effects Models Using lme4. *Journal of Statistical Software* 67, 1–48.
- Belfi, A.M., Karlan, B., and Tranel, D. (2016). Music evokes vivid autobiographical memories. *Memory* 24, 979–989.
- Bisby, J.A., Horner, A.J., Hørlyck, L.D., and Burgess, N. (2016). Opposing effects of negative emotion on amygdalar and hippocampal memory for items and associations. *Social Cognitive and Affective Neuroscience* 11, 981–990.
- Blais-Rochette, C., and Miranda, D. (2016). Music-evoked autobiographical memories, emotion regulation, time perspective, and mental health. *Musicae Scientiae* 20, 26–52.
- Brédart, S., and Barsics, C. (2012). Recalling Semantic and Episodic Information From Faces and Voices: A Face Advantage. *Current Directions in Psychological Science* 21, 378–381.
- Burke, A., Heuer, F., and Reisberg, D. (1992). Remembering emotional events. *Memory & Cognition* 20, 277–290.
- Cady, E.T., Harris, R.J., and Knappenberger, J.B. (2008). Using music to cue autobiographical memories of different lifetime periods. *Psychology of Music* 36, 157–177.
- Cheke, L.G., and Clayton, N.S. (2013). Do different tests of episodic memory produce consistent results in human adults? *Learn. Mem.* 20, 491–498.
- Clayton, N.S., and Dickinson, A. (1998). Episodic-like memory during cache recovery by scrub jays. *Nature* 395, 272–274.
- Cook, S., Kokmotou, K., Soto, V., Fallon, N., Tyson-Carr, J., Thomas, A., Giesbrecht, T., Field, M., and Stancak, A. (2017). Pleasant and unpleasant odour-face combinations influence face and odour perception: An event-related potential study. *Behavioural Brain Research* 333, 304–313.
- Cornell Karnekull, S., Jonsson, F.U., Willander, J., Sikstrom, S., and Larsson, M. (2015). Long-Term Memory for Odors: Influences of Familiarity and Identification Across 64 Days. *Chemical Senses* 40, 259–267.
- Craik, F.I.M., and Lockhart, R.S. (1972). Levels of processing: A framework for memory research. *Journal of Verbal Learning and Verbal Behavior* 11, 671–684.
- Craik, F.I.M., and Tulving, E. (1975). Depth of processing and the retention of words in episodic memory. *Journal of Experimental Psychology: General* 104, 268–294.
- Crookes, K., Ewing, L., Gildenhuis, J., Kloth, N., Hayward, W.G., Oxner, M., Pond, S., and Rhodes, G. (2015). How Well Do Computer-Generated Faces Tap Face Expertise? *PLOS ONE* 10, e0141353.

- Davis, R.G. (1977). Acquisition and retention of verbal associations to olfactory and abstract visual stimuli of varying similarity. *J. Exp. Psychol.-Hum. L.* 3, 37–51.
- Delplanque, S., Grandjean, D., Chrea, C., Aymard, L., Cayeux, I., Le Calvé, B., Velazco, M.I., Scherer, K.R., and Sander, D. (2008). Emotional Processing of Odors: Evidence for a Nonlinear Relation between Pleasantness and Familiarity Evaluations. *Chemical Senses* 33, 469–479.
- Distel, H., Ayabe-Kanamura, S., Martínez-Gómez, M., Schicker, I., Kobayakawa, T., Saito, S., and Hudson, R. (1999). Perception of everyday odors--correlation between intensity, familiarity and strength of hedonic judgement. *Chem Senses* 24, 191–199.
- Doerksen, S., and Shimamura, A.P. (2001). Source memory enhancement for emotional words. *Emotion* 1, 5–11.
- Du, S., Tao, Y., and Martinez, A.M. (2014). Compound facial expressions of emotion. *PNAS* 111, E1454–E1462.
- Easton, A., and Eacott, M.J. (2008a). A new working definition of episodic memory: replacing “when” with “which.” In *Handbook of Episodic Memory*, E. Dere, A. Easton, L. Nadel, and J.P. Huston, eds. (Düsseldorf: Elsevier), pp. 185–196.
- Easton, A., and Eacott, M.J. (2008b). Chapter 2.3 A new working definition of episodic memory: replacing “when” with “which.” In *Handbook of Behavioral Neuroscience*, (Elsevier), pp. 185–196.
- Ernst, A., Bertrand, J.M.F., Voltzenlogel, V., Souchay, C., and Moulin, C.J.A. (2021). The Proust Machine: What a Public Science Event Tells Us About Autobiographical Memory and the Five Senses. *Front. Psychol.* 11.
- Fawcett, J.M., and Ozubko, J.D. (2016). Familiarity, but not Recollection, Supports the Between-Subject Production Effect in Recognition Memory. *Can J Exp Psychol* 70, 99–115.
- Ferreri, L., and Rodriguez-Fornells, A. (2017). Music-related reward responses predict episodic memory performance. *Experimental Brain Research*.
- Ferreri, L., Mas-Herrero, E., Cardona, G., Zatorre, R.J., Antonijoan, R.M., Valle, M., Riba, J., Ripollés, P., and Rodriguez-Fornells, A. (2021). Dopamine modulations of reward-driven music memory consolidation. *Ann N Y Acad Sci*.
- Fuentes-Sánchez, N., Pastor, R., Escrig, M.A., Elipe-Miravet, M., and Pastor, M.C. (2021). Emotion elicitation during music listening: Subjective self-reports, facial expression, and autonomic reactivity. *Psychophysiology* 58, e13884.
- Groussard, M., Chan, T.G., Coppalle, R., and Platel, H. (2019). Preservation of Musical Memory Throughout the Progression of Alzheimer’s Disease? Toward a Reconciliation of Theoretical, Clinical, and Neuroimaging Evidence. *JAD* 68, 857–883.
- Hackländer, R.P.M., Janssen, S.M.J., and Bermeitinger, C. (2018). An in-depth review of the methods, findings, and theories associated with odor-evoked autobiographical memory. *Psychonomic Bulletin & Review*.
- Halpern, A., Talarico, J., Gouda, N., and J. Williamson, V. (2018). Are Musical Autobiographical Memories Special? It Ain’t Necessarily So. *Music Perception* 35, 561–572.
- Hartig, F. (2021). DHARMA: Residual Diagnostics for Hierarchical (Multi-Level / Mixed) Regression Models.
- Herz, R.S. (1998). Are odors the best cues to memory? A cross-modal comparison of associative memory stimuli. *Ann. N. Y. Acad. Sci.* 855, 670–674.

- Herz, R., and Cupchik, G. (1995). The Emotional Distinctiveness of Odor-evoked Memories. *Chemical Senses* 20, 517–528.
- Herz, R., and Engen, T. (1996). Odor memory: Review and analysis. *Psychonomic Bulletin & Review* 3, 300–313.
- Hutmacher, F. (2021). Do you remember? Similarities and differences between the earliest childhood memories for the five senses. *Memory* 29, 345–352.
- Jakubowski, K., Belfi, A., and Eerola, T. (2021a). Phenomenological Differences in Music- and Television-Evoked Autobiographical Memories. *Music Perception* 38, 435–455.
- Jakubowski, K., Belfi, A., and Eerola, T. (2021b). Phenomenological Differences in Music- and Television-Evoked Autobiographical Memories. *Music Perception* 38, 435–455.
- Janata, P., Tomic, S.T., and Rakowski, S.K. (2007). Characterisation of music-evoked autobiographical memories. *Memory* 15, 845–860.
- Jiang, T., Soussignan, R., Rigaud, D., Martin, S., Royet, J.P., Brondel, L., and Schaal, B. (2008). Alliesthesia to food cues: heterogeneity across stimuli and sensory modalities. *Physiol Behav* 95, 464–470.
- Jiang, T., Soussignan, R., Schaal, B., and Royet, J.-P. (2015). Reward for food odors: an fMRI study of liking and wanting as a function of metabolic state and BMI. *Soc Cogn Affect Neurosci* 10, 561–568.
- Johansson, M., Mecklinger, A., and Treese, A.-C. (2004). Recognition memory for emotional and neutral faces: An event-related potential study. *Journal of Cognitive Neuroscience* 16, 1840–1853.
- Jraissati, Y., and Deroy, O. (2021). Categorizing Smells: A Localist Approach. *Cognitive Science* 45, e12930.
- Kaplan, R.L., Van Damme, I., Levine, L.J., and Loftus, E.F. (2016). Emotion and False Memory. *Emotion Review* 8, 8–13.
- Kätsyri, J. (2018). Those Virtual People all Look the Same to me: Computer-Rendered Faces Elicit a Higher False Alarm Rate Than Real Human Faces in a Recognition Memory Task. *Front Psychol* 9.
- Keightley, M.L., Chiew, K.S., Anderson, J.A.E., and Grady, C.L. (2011). Neural correlates of recognition memory for emotional faces and scenes. *Soc Cogn Affect Neurosci* 6, 24–37.
- Kensinger, E.A., and Corkin, S. (2003). Memory enhancement for emotional words: Are emotional words more vividly remembered than neutral words? *Memory & Cognition* 31, 1169–1180.
- Kuznetsova, A., Brockhoff, P., and Christensen, R. (2015). *LmerTest: Tests in linear mixed effects models. R Package Version 2.*
- Larsson, M., Öberg, C., and Bäckman, L. (2006). Recollective experience in odor recognition: Influences of adult age and familiarity. *Psychological Research* 70, 68–75.
- Larsson, M., Blåvarg, E.C., and Jönsson, F. (2009). Bad Odors Stick Better Than Good Ones: Olfactory Qualities and Odor Recognition. *Experimental Psychology* 56, 375–380.
- Larsson, M., Willander, J., Karlsson, K., and Arshamian, A. (2014). Olfactory LOVER: behavioral and neural correlates of autobiographical odor memory. *Front. Psychol.* 5.
- Lenth, R.V., Buerkner, P., Herve, M., Love, J., Riebl, H., and Singmann, H. (2020). *emmeans: Estimated Marginal Means, aka Least-Squares Means.*

- Levine, L.J., and Pizarro, D.A. (2004). Emotion and Memory Research: A Grumpy Overview. *Social Cognition* 22, 530–554.
- Levy, D.A. (2004). Impaired odor recognition memory in patients with hippocampal lesions. *Learning & Memory* 11, 794–796.
- Liu, C.H., Chen, W., and Ward, J. (2014). Remembering faces with emotional expressions. *Frontiers in Psychology* 5, 1439.
- Lockhart, R.S., and Murdock, B.B. (1970). Memory and the theory of signal detection. *Psychol. Bull.* 74, 100–109.
- Lüdecke, D., Ben-Shachar, M., Patil, I., Waggoner, P., and Makowski, D. (2021). performance: An R Package for Assessment, Comparison and Testing of Statistical Models. *JOSS* 6, 3139.
- MacDorman, K.F., Green, R.D., Ho, C.-C., and Koch, C.T. (2009). Too real for comfort? Uncanny responses to computer generated faces. *Computers in Human Behavior* 25, 695–710.
- Majid, A., and Burenhult, N. (2014). Odors are expressible in language, as long as you speak the right language. *Cognition* 130, 266–270.
- Makowski, D., Ben-Shachar, M., and Lüdecke, D. (2020). The easystats collection of R packages. (easystats).
- Mandaïron, N., Poncelet, J., Bensafi, M., and Didier, A. (2009). Humans and Mice Express Similar Olfactory Preferences. *PLoS One* 4, e4209.
- Mantel, M., Ferdenzi, C., Roy, J.-M., and 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 Topogr* 32, 977–986.
- Mao, X., You, Y., Li, W., and Guo, C. (2015). Emotion impairs extrinsic source memory—An ERP study. *Biological Psychology* 110, 182–189.
- Maquet, P. (2001). The role of sleep in learning and memory. *Science* 294, 1048–1052.
- Mas-Herrero, E., Marco-Pallarés, J., Lorenzo-Seva, U., Zatorre, R., and Rodriguez-Fornells, A. (2013). Individual Differences in Music Reward Experiences. *Music Perception: An Interdisciplinary Journal* 31, 118–138.
- Mas-Herrero, E., Zatorre, R.J., Rodriguez-Fornells, A., and Marco-Pallarés, J. (2014). Dissociation between Musical and Monetary Reward Responses in Specific Musical Anhedonia. *Current Biology* 24, 699–704.
- Matuschek, H., Kliegl, R., Vasishth, S., Baayen, H., and Bates, D. (2017). Balancing Type I error and power in linear mixed models. *Journal of Memory and Language* 94, 305–315.
- Meissner, C.A., and Brigham, J.C. (2001). Thirty years of investigating the own-race bias in memory for faces: A meta-analytic review. *Psychology, Public Policy, and Law* 7, 3–35.
- Meng, X., Zhang, L., Liu, W., Ding, X., Li, H., Yang, J., and Yuan, J. (2017). The impact of emotion intensity on recognition memory: Valence polarity matters. *International Journal of Psychophysiology* 116, 16–25.
- Olofsson, J.K., and Gottfried, J.A. (2015). The muted sense: neurocognitive limitations of olfactory language. *Trends in Cognitive Sciences* 19, 314–321.
- Palombo, D.J., Sheldon, S., and Levine, B. (2018). Individual Differences in Autobiographical Memory. *TRENDS COGN. SCI.* 22, 583–597.

- Plailly, J., Luangraj, N., Nicklaus, S., Issanchou, S., Royet, J.P., and Sulmont-Rosse, C. (2011). Alliesthesia is greater for odors of fatty foods than of non-fat foods. *Appetite* 57, 615–622.
- Plailly, J., Villalba, M., Vallat, R., Nicolas, A., and Ruby, P. (2019). Incorporation of fragmented visuo-olfactory episodic memory into dreams and its association with memory performance. *Sci Rep* 9, 15687.
- Rey, L., Désoche, C., Thévenet, M., Garcia, S., Tillmann, B., and Plailly, J. (a). Episodic memory for odors, music and faces – an ecological laboratory approach using virtual reality.
- Saive, A.-L., Ravel, N., Thévenet, M., Royet, J.-P., and Plailly, J. (2013). A novel experimental approach to episodic memory in humans based on the privileged access of odors to memories. *Journal of Neuroscience Methods* 213, 22–31.
- Saive, A.-L., Royet, J.-P., Ravel, N., Thévenet, M., Garcia, S., and Plailly, J. (2014a). A unique memory process modulated by emotion underpins successful odor recognition and episodic retrieval in humans. *Frontiers in Behavioral Neuroscience* 8.
- Saive, A.-L., Royet, J.-P., and Plailly, J. (2014b). A review on the neural bases of episodic odor memory: from laboratory-based to autobiographical approaches. *Frontiers in Behavioral Neuroscience* 8.
- Saive, A.-L., Royet, J.-P., Garcia, S., Thévenet, M., and Plailly, J. (2015). “What-Where-Which” Episodic Retrieval Requires Conscious Recollection and Is Promoted by Semantic Knowledge. *PLOS ONE* 10, e0143767.
- Salakka, I., Pitkäniemi, A., Pentikäinen, E., Mikkonen, K., Saari, P., Toiviainen, P., and Särkämö, T. (2021). What makes music memorable? Relationships between acoustic musical features and music-evoked emotions and memories in older adults. *PLOS ONE* 16, e0251692.
- Schacter, D.L. (1999). *Insights From Psychology and Cognitive Neuroscience*. *American Psychologist* 22.
- Schulkind, M.D., and Woldorf, G.M. (2005). Emotional organization of autobiographical memory. *Memory & Cognition* 33, 1025–1035.
- Sezille, C., Messaoudi, B., Bertrand, A., Jousain, P., Thévenet, M., and Bensafi, M. (2013). A portable experimental apparatus for human olfactory fMRI experiments. *J. Neurosci. Methods* 218, 29–38.
- Sheldon, S., and Donahue, J. (2017). More than a feeling: Emotional cues impact the access and experience of autobiographical memories. *Memory & Cognition* 45, 731–744.
- Sheldon, S., Williams, K., Harrington, S., and Otto, A.R. (2020). Emotional cue effects on accessing and elaborating upon autobiographical memories. *Cognition* 198, 104217.
- Singmann, H., Bolker, B., Westfall, J., Aust, F., Ben-Shachar, M.S., Højsgaard, S., Fox, J., Lawrence, M.A., Mertens, U., Love, J., *et al.* (2021). *afex: Analysis of Factorial Experiments*.
- Smith, S.A. (2019). Virtual reality in episodic memory research: A review. *Psychon Bull Rev* 26, 1213–1237.
- Stalinski, S.M., and Schellenberg, E.G. (2013). Listeners remember music they like. *Journal of Experimental Psychology: Learning, Memory, and Cognition* 39, 700–716.
- Stickgold, R. (2005). Sleep-dependent memory consolidation. *Nature* 437, 1272–1278.
- Tinwell, A., Grimshaw, M., Nabi, D.A., and Williams, A. (2011). Facial expression of emotion and perception of the Uncanny Valley in virtual characters. *Computers in Human Behavior* 27, 741–749.

- Tulving, E. (1972). Episodic and semantic memory. In *Organization of Memory*, (Oxford, England: Academic Press), pp. xiii, 423–xiii, 423.
- Tulving, E. (1983). *Elements of Episodic Memory* (Oxford: Clarendon Press).
- Tulving, E. (1985). Memory and consciousness. *Canadian Psychology/Psychologie Canadienne* 26, 1–12.
- Tulving, E. (2002). Episodic Memory: From Mind to Brain. *Annu. Rev. Psychol.* 53, 1–25.
- Vieillard, S., Peretz, I., Gosselin, N., Khalfa, S., Gagnon, L., and Bouchard, B. (2008). Happy, sad, scary and peaceful musical excerpts for research on emotions. *Cognition and Emotion* 22, 720–752.
- Wang, B. (2013). Facial expression influences recognition memory for faces: Robust enhancement effect of fearful expression. *Memory* 21, 301–314.
- Wright, D.B., and London, K. (2009). Multilevel modelling: Beyond the basic applications. *British Journal of Mathematical and Statistical Psychology* 62, 439–456.
- Wright, D.B., Horry, R., and Skagerberg, E.M. (2009). Functions for traditional and multilevel approaches to signal detection theory. *Behavior Research Methods* 41, 257–267.

5. Supplementary materials

5.1 Encoding variables and their influence on memory performance

5.1.1 Data analyses

The Number of Presentation (*i.e.*, number of clicks on the box) and their day of presentation (Encoding day; day1, day2, day3) were computed. The relationship between the encoding variables and the episodic memory performance was investigated by analyzing the effect of Number of Presentation and Encoding day (day1, day2, day3) on each of the 7 steps of interest of the episodic memory process (from Cue to Cue+, from Cue+ to Room+ and then to Room+Period+, from Cue+ to Period+ and then to Room+Period+, from Cue to Room+Period+, from Cue+ to Room+Period+, from Cue to any episodic retrieval, from Cue+ to any episodic retrieval).

For each step of interest, a model was fitted to predict the probability to succeed, with as fixed effects: Number of Presentation, Encoding Day, and the three second-order interactions between Sensory Modality, Number of Presentation and Encoding Day. The maximal model of fixed effect with 3 Encoding Day (day1, day2, day3) x 3 Sensory Modality (odor, music, face) x 1 Number of Presentation was not fitted since it was not powerful enough. For recognition response step, maximal random effect structure was: (Number of Presentation * Encoding Day + Sensory Modality: Number of Presentation + Sensory Modality: Encoding Day | Participant) + (Number of Presentation * Encoding Day | Item). For other episodic steps, random interactions were not fitted in the maximal random model, as they were too few observations for such a complex model. The maximal random effect structure was: (Number of Presentation + Encoding Day + Sensory Modality | Participant) + (Number of Presentation + Encoding Day | Item).

5.1.2 Results

During encoding, the Number of Presentation, *i.e.*, the amount of time each cue was explored, was of 5.06 ± 0.71 (mean \pm SD) for odor, 5.27 ± 0.59 for music, and 5.14 ± 0.34 for faces.

The influence of Number of Presentation and Encoding Day, *i.e.*, day1, day 2 and day 3 on each step that was above chance level of the episodic memory process was tested.

5.1.2.1 From Cue to Cue +

The model was fitted with [(Encoding day || Participant) + (1 | Item)] as random effects (conditional $R^2 = 0.38$, marginal $R^2 = 0.34$). A significant interaction between Sensory Modality and Encoding Day ($\chi^2_{(6)} = 15.49$, $p < .05$), revealed no primacy or recency effects for odors, music, and faces ($z < 1.99$, p 's > 0.14).

5.1.2.2 From Cue+ to Room+

The model was fitted using [(Number of Presentation||Participant) + (Number of Presentation||Item)] random effect structure (conditional $R^2 = 0.03$ marginal $R^2 = 0.03$). A significant interaction between Sensory Modality and Number of Presentation ($\chi^2_{(2)} = 7.32, p < .05$) revealed that the slope for music (-0.04 ± 0.14) was significantly different than the slope for odor ($-0.08 \pm 0.15; z = 2.72, p < .05$). The Number of Presentation slope was not different from zero for each sensory modality (z 's $< 0.55; p$'s > 0.59).

5.1.2.3 From Cue+ to Period+

The model was fitted using [(1 | Participant) + (1 | Item)] random effect structure (conditional $R^2 = 0.08$, marginal $R^2 = 0.04$). A significant interaction between Sensory Modality and Number of Presentation ($\chi^2_{(2)} = 8.70, p < .05$) revealed that the slope for music (0.03 ± 0.15) was significantly different than the slope for odor ($0.13 \pm 0.15, z = 2.71, p < .05$) and tended to be different than the slope for face ($0.12 \pm 0.15, z = 2.38, p = 0.052$). However, Number of Presentation slope for each sensory modality was not different from zero (z 's $< 0.89, p$'s $> .38$).

5.1.2.4 From Cue to Room+Period+

The model was fitted using [(1|Participant) + (1|Item)] random effect structure (conditional $R^2 = 0.19$, marginal $R^2 = 0.08$). It revealed a significant principal effect of Number of Presentation (slope = $0.40 \pm 0.18, \chi^2_{(1)} = 4.47, p < .05$). A significant interaction between Sensory Modality and Encoding Day ($\chi^2_{(4)} = 11.16, p < .05$), revealed no clear primacy or recency effects for odors, music, and faces (z 's $< 2.14, p$'s > 0.09), except for two comparisons showing that faces evoked more retrieval the first day as compared to the third day ($z = 2.53, p < .05$), and that odors evoked more retrieval the the third day as compared to the second day ($z = 2.42; , p < .05$).

5.1.2.5 From Cue+ to Room+Period+

The model was fitted using [(1|Participant)] random effect structure (conditional $R^2 = 0.19$, marginal $R^2 = 0.09$) with the *Bobyqa* optimizer. A significant interaction between Sensory Modality and Number of Presentation ($\chi^2_{(2)} = 8.13, p < .05$) revealed that the slope for odor (0.46 ± 0.19) was significantly different than the slope for music ($0.31 \pm 0.19, z = 2.80, p < .05$). Moreover, Number of Presentation slope was different from zero only for odor ($z = 2.37, p < .05$) and face ($0.41 \pm 0.20, z = 2.09, p < .05$), but not music ($z = 1.67, p = .10$) cues.

5.1.2.6 From Cue to any episodic retrieval

The model was fitted using [(1 | Participant)] random effect structure (conditional $R^2 = 0.04$, marginal $R^2 = 0.03$). A significant interaction between Sensory Modality and Number of

Presentation ($\chi^2_{(2)} = 6.00, p < .05$) revealed that the slope for face (-0.09 ± 0.13) was significantly different than the slope for odor ($-0.01 \pm 0.13, z = 2.42, p < .05$). The Number of Presentation slope for each Sensory Modality was not different from zero (z 's < 0.69 ; p 's > 0.49). A significant interaction between Sensory Modality and Encoding Day ($\chi^2_{(4)} = 12.64, p < .05$), revealed a recency effect for odors which evoked more retrieval the third day as compared to the first day ($z = 2.82$; , $p < .05$). No clear primacy or recency effects were shown for music, and faces (z 's $< 2.11, p$'s $> .11$)

5.1.2.7 From Cue+ to any episodic retrieval

The model was fitted using [(1 | Participant) + (Number of Presentation || Item)] random effect structure (conditional $R^2 = 0.06$, marginal $R^2 = 0.05$). A significant interaction between Sensory Modality and Number of Presentation ($\chi^2_{(2)} = 13.16, p < .01$) revealed that the slope for music (-0.18 ± 0.14) was significantly different than the slope for odor ($-0.03 \pm 0.15, z = 3.76, p < .001$). The Number of Presentation slope for each Sensory Modality was not different from zero (z 's < 1.28 ; p 's > 0.20).

Study 3: Neural basis of episodic memory for emotional odors and music

This study was conducted with the collaboration of Philippe Litaudon, Gaëlle Leroux, Clément Désoche, Marc Thévenet, Samuel Garcia, Christelle Daudé, Barbara Tillmann, and Jane Plailly. LR, GL, BT and JP designed the experiment. CID created the virtual reality software. MT created the olfactometer and the software for episodic memory retrieval testing. LR, PL, ChD and JP ran the experiment. SG extracted and gathered the raw data. LR, PL, GL, BT and JP analyzed and interpreted the data. LR, JP and BT wrote the manuscript.

Thirty participants were intended to be included in this experiment. A group of 9 was included before the first COVID19 lockdown, and a tenth participant was included very recently. Given the thesis program, it was impossible to include more participants for this manuscript. Thus, the following part report pilot results obtained in a group of 10 participants.

1 Introduction

Olfaction is presented as a privileged gateway to episodic memories (for reviews, Annett, 1996; Hackländer et al., 2018; Herz and Engen, 1996b; Larsson et al., 2014; Saive et al., 2014b). This hypothesis has been investigated by comparing voluntary evoked memories of participants' life events that are triggered by odors cues to those evoked by images, sounds and labels (de Bruijn and Bender, 2017; Chu and Downes, 2002; Ernst et al., 2021; Goddard et al., 2005; Herz, 2004; Herz and Schooler, 2002; Karlsson et al., 2013; Miles and Berntsen, 2011; Rubin et al., 1984; Willander et al., 2015). These studies have congruently shown that, while olfactory cues evoke less autobiographical memories, the evoked memories are more emotional, older, and less often rehearsed, thought of and talked about than memories evoked by other sensory modality cues. The common hypothesis that explains the singularity of odor-evoked memories suggests that there are close and direct anatomical links between the primary olfactory brain regions, memory and emotional structures (Daniels and Vermetten, 2016; Saive et al., 2014b). The odor signal coming from olfactory receptor neurons arrives at the third synapse to the hippocampus through the entorhinal cortex, areas that has been related to odor associative memory (Baudry, 2020; Saive et al., 2014b). It also takes only three synapses for an odor signal to reach the basolateral amygdala, that is involved in the relationship between odor perception and emotion (Kontaris et al., 2020; Royet et al., 2003; Yeshurun and Sobel, 2010). Amygdala is also functionally connected with areas that are involved in memory, including hippocampus and entorhinal cortex (Arnold et al., 2020). And, in contrast with other sensory modalities, projections from the sensory input onto limbic areas do not pass via the thalamus (Cahill et al., 1995; Phelps and Anderson, 1997). Functionally, odor-evoked memories are associated with greater activation of regions involved in the

recollection of autobiographical memories and in emotional processing, than memories evoked by verbal or visual formats of the odor objects (Arshamian et al., 2013; Herz, 2004).

However, episodic and associative memories evoked by odors are still understudied and they have almost never been compared to memory evoked by cues that also share close links with emotion and memory (Davis, 1977b; Herz, 1998; Herz and Cupchik, 1995b). Recently, we ran two ecological-laboratory tasks comparing episodic memory evoked by odors, to episodic memory evoked by music or faces (Rey et al., b, a), which are stimuli with strong ecological validity, and which demonstrate specificities for evoking episodic memories (Belfi et al., 2016; Brédart and Barsics, 2012b). Our findings support the hypothesis that odor-evoked episodic memories are influenced by the emotion associated to the odor. They provided evidence for a remembering pattern that differs according to the sensory modality of the episodic memory cue and its emotion. Notably, we found two main results: 1) Music was better recognized than associated to other episode dimensions, while the reverse was observed for odors, and 2) Emotion favored enhanced recognition for music stimuli and enhanced recollection for odor stimuli. Our results pointed out that music and odors may both rely on distinct processes, while both being presented as good episodic and autobiographical memories cues that are both highly emotional. Music is indeed reported to be consistently associated with strong emotions (Janata et al., 2007; Schulkind et al., 1999), and music-evoked autobiographical memories contain more perceptual details (Belfi et al., 2016) and were described as being more vivid, more significant, more pleasant and emotionally more intense than face stimuli or TV shows (Belfi et al., 2016; Jakubowski et al., 2021). The recall of music-evoked AM activates a majority of brain regions as those evoked by other non-musical AM: Posterior cingular cortex, lateral temporal lobes, parahippocampal gyrus, medial prefrontal cortex, with the exception of the medial temporal lobe which was not observed in all studies investigating music-evoked AM (Ford et al., 2011; Janata, 2009a).

Research is currently still insufficient to sustain the hypothesis to explain the singularity of odor-evoked memories by the limbic nature of the olfactory pathway, notably because odor cues were insufficiently compared with other emotional cues. The present study aims to bridge this gap by identifying the neural networks that are modulated by the subjective recollection associated with episodic recall, as well as the sensory modality and emotion (negative, positive) of the recall cue. To this aim, participants discovered three complex episodes in three days in semi-immersive virtual reality, where devices diffusing an odor or a music stimulus were disposed. On the fourth day, their memory of these episode was tested during a functional magnetic resonance imaging (fMRI) recording. The acquired functional images will then be analyzed with an effective connectivity technique, referred to as dynamic causal modeling,

aiming to study the effect of sensory modality and emotion on the network that links primary sensory areas identified with localizer runs, to the amygdala and the hippocampus. We made the hypothesis that this network will be more strongly activated for the olfactory cue condition than the music cue condition.

2 Materials and Methods

2.1 Ethics Statement

All participants provided written informed consent prior to participation. The protocol was approved by the national Institutional Review Board, according to French regulations for biomedical experiments with healthy volunteers [Ethical Committee of CPP IDF8 (April 25,2017), ID RCB: 2016-A01931-50].

2.2 Participants

Ten healthy participants [8 women; aged 23.1 ± 4.5 (mean \pm standard deviation)] consented to participate in this experiment and received 120 Euros in compensation. All participants reported normal senses of smell and no visual or auditory impairments. They were recruited through campus electronic mails and posters, and social networks. Two participants were discarded due to technical failure during the fMRI session, due to software issues on the devices that sent odorants and controlled the trackball.

2.3 Stimuli and materials

2.3.1 Multidimensional episodes

A software called EpisOdor-Version3 was developed using Unity 5.6.5 (Unity Technologies, USA) by a local technical platform *NeuroImmersion* (Cesame Institute, CRNL). EpisOdor allowed the presentation of a virtual 3D house in a first-person view (non-immersive virtual reality) that can be actively explored using an amagnetic trackball (Nata Technologies, Coquitlam, BC, Canada) (Figure1A). This house was presented in three conditions created with variation of light intensity and color: day, night, and twilight. A bright environment corresponded to daylight, a dark environment corresponded to nightlight, and an intermediate orangish light corresponded to twilight (defining the *Which context* component). The house was composed of three rooms (defining the *Where* component), a bedroom, an office and a living room, connected to a corridor by closable doors. Each of these rooms was defined by its furniture (e.g., chair, bed, table, piano) and decorative elements (e.g., painting, carpet). In addition, two clickable objects, a perfume diffuser and a radio set, were placed in each room. In two of the three rooms these objects delivered respectively an odor and a musical stimulus (defining the *What* component), in the third room, these objects did not deliver anything, they were non-clickable. The clickable delivering objects were highlighted by an arrow that appeared when participants were nearby. Objects were relatively distantly located from each

other within a given room and were arranged differently for each room. To avoid any emotional contagion between stimuli in the same room (Pützer et al., 2019), the two rooms with clickable objects were assigned to stimuli of one emotional type each (negative or positive). In each period, a room presented emotionally negative objects, another room presented emotionally positive objects and the last room presented non-clickable objects. This repartition changed pseudo randomly among the three days so that each room had clickable objects twice (once emotionally negative, once emotionally positive) and a non-clickable object on one day.

2.3.2 Stimuli and apparatus

For the main experimental task, six stimuli of each sensory modality (Odor, Music; leading to a total of 12 stimuli) were selected. In each sensory modality, stimuli set was composed of three emotionally negative stimuli and three emotionally positive stimuli. Stimuli were selected from the previous experiment (Rey et al., b). The selection was based on the valence, emotional intensity, and familiarity of the stimuli, in order to favor the comparability of the stimuli between sensory modalities. The three negative and the three positive stimuli of each sensory modality were pseudo-randomly assigned to a room and a period for each participant.

2.3.2.1 Odorants

Odorants consisted of essential oils, single or mixtures of monomolecular chemical compounds, and fragrances. Negative odorants were composed of Musk (Givaudan-Roure), Stemone (Créations aromatiques), and Vetiver (Davenne, Montfavet, France). Positive odorants were composed of Bien-Être (EmoSens), Osmose (EmoSens), and Ylang Ylang (Givaudan). The undiluted odorants were placed in 10-ml U-shaped Pyrex® tubes (VS Technologies, Saint-Priest, France) filled with Pebax® microporous substances (Arkema, Colombes, France).

The odorants were presented with a 20-channel computer-controlled olfactometer adapted from (Sezille et al., 2013). This odor diffusion system was developed to synchronize odorous stimuli with breathing. The participants' nasal respiratory signals were acquired using a nasal canula and were used to trigger the odor stimulation through an airflow sensor. During odor stimulation, the olfactometer waited for the participants' subsequent expiration, allowing the odor to be perceived at the beginning of the following inspiration. When this expiration was detected, an unodorized airflow was sent to one of the U-shaped odorous tubes. Odorized airflow and air carrier were sent to and mixed in a homemade mixing head made of polytetrafluoroethylene. It was connected to the nostrils through two Teflon tubes, fixed to the nasal canula, opening out under the nostrils. The olfactometer airflow rate was set between 3.0 and 4.5 l/min depending on the odorant physical intensity. The odorants were delivered over a period of 6 s. The olfactometer was controlled by an in-house LabView software

(National Instruments, Austin, TX, USA) and interacted with EpisOdor-Version3 software to send an odor after a click on the perfume diffuser. The olfactometer was conceived to be compatible with magnetic fields of the fMRI, the magnetic part being in the control room and the amagnetic part being in the scanner room, as close as possible to the nose of the participants.

2.3.2.2 Musical Pieces

Musical clips were selected from our previous experiment (Rey et al., b), in which we adapted the materials used in Vieillard et al. (2008) (Copyright, Bernard Bouchard, 1998) (http://www.peretzlab.ca/knowledge_transfer/). They were modified in MIDI (Digital Performer®, MOTU, Cambridge, USA). The musical clips were played with an acoustic piano timbre (Cubase®, Steinberg Media Technologies, Hamburg Germany) and presented with EpisOdor-Version3 and our in-house LabView software at a comfortable loudness level with headphones. Their average duration of the excerpts was 6.62 ± 1.12 s.

In the retrieval task, in order to make odor and music presentations as similar as possible, not only odors but also the musical excerpts were presented according to the participants' breathing cycles. In the experimental room, regular headphones covering ears were used. When in the fMRI scanner, amagnetic OptoActive™ active noise cancelling headphones (Optoacoustics Ltd., Mazor, Israel) were used.

2.4 Experimental procedure

The main experiment was divided into four sessions (Figure 1B), one session per day on consecutive days. The first three days were devoted to the encoding phase and took place in an experimental room of the laboratory, and the last day was devoted to memory assessment and the brain data recording and took place in the fMRI scanner of the CERMEP (Bron, France). A night of sleep followed each of the encoding sessions to promote consolidation and to reduce interference (Abichou et al., 2019; Maquet, 2001; Stickgold, 2005). Participants completed the four sessions at the same time of the day to limit the differential influences of internal states (hunger, satiety) on olfactory and cognitive processes between sessions (Jiang et al., 2008; Plailly et al., 2011). The participants were requested to breathe through their nose as normally as possible without consciously modifying their respiration.

On the fourth day, three additional sessions took place. Two of them (i.e., localizer tasks) were done in the fMRI scanner, and the third one (i.e., a stimuli sensory evaluations task) was run outside the fMRI scanner in the control room.

2.4.1 Encoding

Each encoding session consisted of at least 10 min during which participants freely explored the house presented in a different period (day, night, and twilight) each of the three days. The order of the period across days was counterbalanced across participants. Participants were encouraged to pay attention to each detail of the rooms and were told to click at least five times on each clickable object to be sure they sufficiently explored each stimulus. The fifth click was indicated to the participant with a change of the box's arrow color, from green to red. No instructions for memorization were given, but participants were informed that they would be questioned about their perception of the episodes on the fourth day. This procedure aimed to ensure free encoding like in real-life situations. Importantly, given the implicit nature of the experiment, they were told that the study aimed to investigate the perception of various environments involving music, and odors. To note, due to software issues, some clicks on the perfume diffuser were not considered by the olfactometer. Some participants thus explored odors less than 5 times.

2.4.2 Retrieval

The retrieval test was displayed with an in-house LabView software. It was composed of three runs organized similarly (Figure 1B). Each run was composed of 24 trials divided in two sessions, each devoted to a sensory modality (odor, music; Figure 1C). The order of the sessions was counterbalanced between runs and participants. Each session began with 15 s of rest introducing the sensory modality of the session ("*Odor*", "*Music*"). Each trial started with 4 s of rest with a fixation cross presented in the middle of the screen (Figure 1D). Then a unimodal cue (Odor or Music) was presented at the next expiration. Each cue was presented two times per run and pseudo-randomly, with the constraint to avoid the presentation of two same odors or music in a row. Participants were first asked whether they subjectively remember the memory associated to the cue (the room and the period; Subjective recollection response: "*Does this cue evoke a memory? (Y/N)*"), allowing metacognitive knowledge evaluation. The answer was given with right ("Yes") or left ("No") click on the trackball. In case of a "No" response, fixation cross appears for 5 s. In case of a "Yes" response, a screen with the instruction "*Immerse yourself into this memory*" was displayed for 5 s. Then, a screen with 6 images arranged in 2 rows x 3 columns appeared. The 3 images in the upper side of the screen represented 3 pictures of the house taken from an aerial view, with respectively the day, twilight, and night contexts. These pictures were placed above the question "*In which context?*". The 3 images in the lower side represented 3 pictures of the rooms that were turned into black and white, with respectively the office, the living room, and the bedroom. These pictures were placed above the question "*In which room?*". Participants had to click on one period and one room picture, in any order, to end the trial (Episodic response). Two seconds

of delay had been added after the second answers to allow participants to eventually change their mind. In total, each cue was presented 6 times during the whole experiment, leading to 18 trials per Emotional Category for each Sensory Modality.

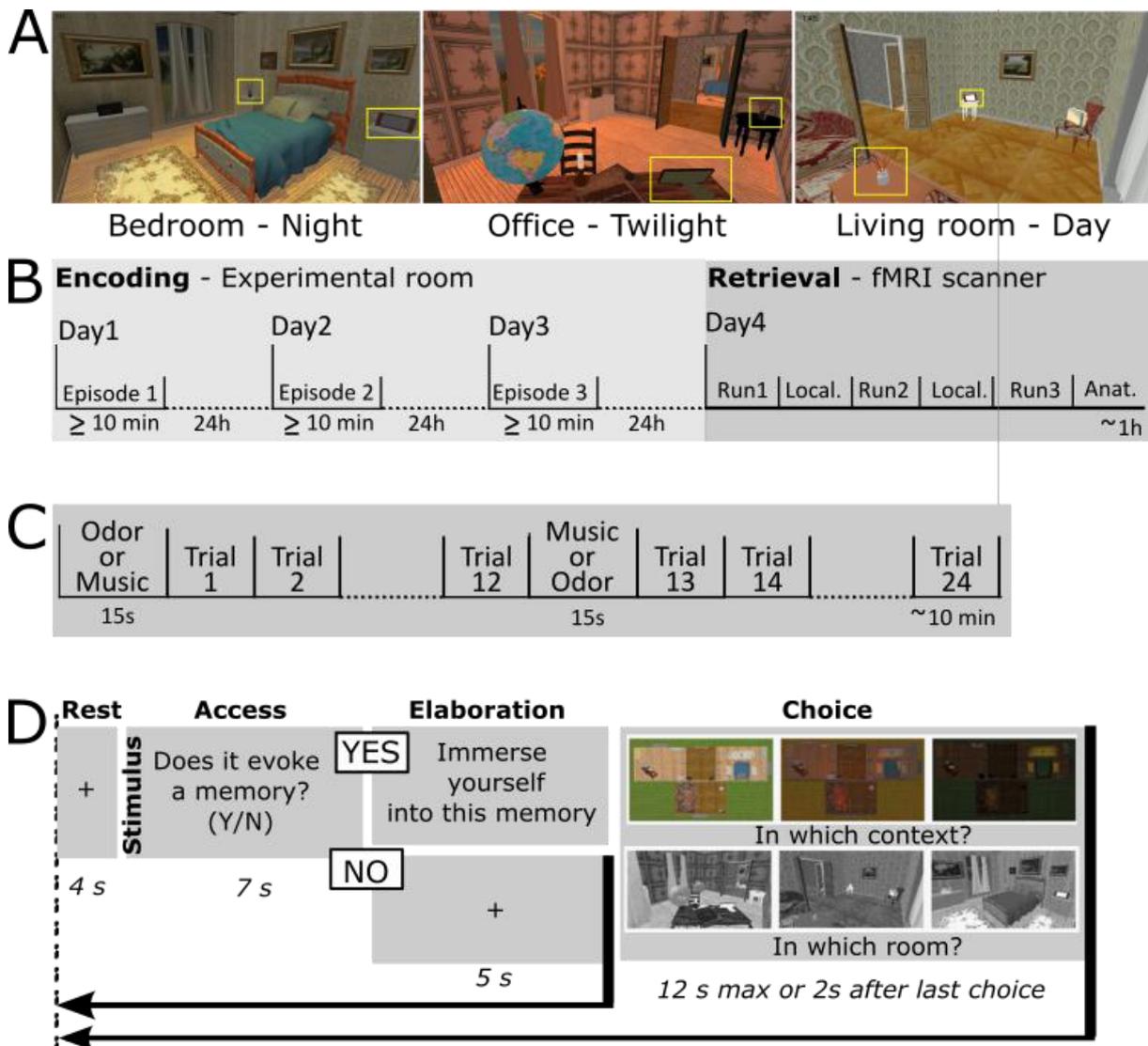


Figure 18. Experimental protocol (A) The three rooms of the virtual house in three different periods of the day (Bedroom in nightlight, Office in twilight, Living-room in daylight). Yellow rectangles represent the clickable objects with which the stimuli were presented. (B) The temporal course of the encoding and the retrieval sessions. During the encoding, participants discovered one episode a day over three days. The fourth day takes place in the fMRI scanner, where the memory of the episodes was tested. (C) Example of a run of the main experiment. It was divided into two parts, one dedicated to each sensory modality. After 15s of instructions, twelve trials followed. (D) Example of a trial of the retrieval phase. After a short period of rest, a stimulus was sent to the participants (Access), which were required to indicate their subjective recollection. A 'No' response was followed by 5 s of rest. A 'Yes' response was followed by 5 s of re-evocation (Elaboration). Then, participants selected a period of time (context) and a room (location) in which they thought this cue was presented at encoding (Choice) and the trial ended.

2.4.3 Sensory evaluations

At the end of the experiment, in the control room, participants were required to rate the 12 stimuli in terms of pleasantness (unpleasant – neutral – pleasant), emotional intensity (very weak – very intense), familiarity (unknown – very familiar), physical intensity (extremely weak – extremely strong), and describability (*i.e.*, capacity to describe the stimulus with a few words; hardly – easily) using non-graduated scales. The pleasantness scale was divided into two equal parts, with a “neutral” value separating the ratings of unpleasantness and pleasantness. The odors were presented with the olfactometer, similarly to the experimental conditions of the main experiment.

2.5 Behavioral data analysis

2.5.1 Subjective recollection

Subjective recollection responses were analyzed through the proportion of “Yes” responses (*vs.* Yes+No responses). It was analyzed as a function of Sensory Modality (music, odor), and Emotional Category (negative, positive) of the cues.

2.5.2 Memory performance

Four response categories were defined based on the room and period selected: *Room-Period-* when both inaccurate room and period were selected (completely incorrect episodic response), *Room-Period+* when the inaccurate room but accurate period were selected and *Room+Period-* when the accurate room but inaccurate period were selected (both partially correct episodic response), and *Room+Period+* when the participants selected both the accurate room and period (corresponding to a complete correct episodic response). The proportions of each episodic response were analyzed as a function of the Sensory Modality (music, odor) and the Emotional Category (negative, positive). They were also compared to chance, with 0.44 for *Room-Period-* [2 rooms out of 3 * 2 periods out of 3], 0.22 for *Room+Period-* and *Room-Period+* [1 dimension out of 3 * 2 dimensions out of 3], and 0.011 for *Room+Period+* [1 room out of 3 * 1 period out of 3].

2.5.3 Stimuli sensory evaluation

Pleasantness, emotional intensity, familiarity, physical intensity and describability ratings of the stimuli were *a posteriori* transformed into scores from 0 to 10. They were Z-scored $[(x-\mu)/\sigma]$ at the individual level to remove bias due to inter-individual differences. They were analyzed as a function of Sensory Modality (odor, music) and Emotional Category (negative, positive).

1.1 Statistical analyses

Statistical analyses were performed with JASP software (JASP Team, 2020). Results were analyzed with one-samples *t*-tests (two-tailed) and repeated measures ANOVAs that were followed by *posthoc* two-sample Student *t*-tests corrected by Bonferroni when main effects

and/or interactions were significant. Results were considered significant at $p < .05$., and a p -value between 0.05 and 0.09 was referred to as a tendency or marginal significance. Their values were corrected with Greenhouse-Geisser when sphericity tests were found significant. When the interaction between two factors was statistically significant, the effects of the main factors were reported, but not further interpreted. Data were described with the means and the standard error of the mean (mean \pm SEM).

2.6 fMRI data analysis

2.6.1 fMRI data acquisition

Functional imaging was conducted using Siemens Prisma 3T scanner (Siemens Medical®, Erlangen, Germany) to acquire gradient-echo T2*-weighted echoplanar images (EPIs) with blood oxygen level-dependent (BOLD) contrast, using a 64-Channel Head/Neck coil. Three runs, of approximately 350 volumes each, were collected in an interleaved ascending sequence (20 slices per volume the whole brain). Imaging parameters were as follows: echo time (TE), 30 ms; repetition time (TR), 1500 ms; slice thickness, 2.3 mm; in-plane resolution, 1.767 * 1.767 mm; field of view, 220 * 220 mm; matrix size, 96 * 96 voxels. Whole-brain high-resolution T1-weighted anatomical scans (0.8 mm³) were acquired after functional scanning (matrix size, 280 * 280; TE, 19.3 ms; TR, 3000 ms; TI, 1000ms; flip angle, 8°). Visual stimuli were projected on a mirror located in front of participants' eyes from a video projector located in the back of the scanner.

2.6.2 fMRI data analysis

Results included in this draft come from preprocessing performed using fMRIPrep 20.2.1 ((Esteban et al., 2021a, 2021b) RRID:SCR_016216), which is based on Nipype 1.5.1 ((Gorgolewski et al., 2011) RRID:SCR_002502).

2.6.2.1 Anatomical data preprocessing

A total of 1 T1-weighted (T1w) images were found within the input BIDS dataset. The T1-weighted (T1w) image was corrected for intensity non-uniformity (INU) with N4BiasFieldCorrection (Tustison et al., 2010), distributed with ANTs 2.3.3 ((Avants et al., 2008)RRID:SCR_004757), and used as T1w-reference throughout the workflow. The T1w-reference was then skull-stripped with a Nipype implementation of the antsBrainExtraction.sh workflow (from ANTs), using OASIS30ANTs as target template. Brain tissue segmentation of cerebrospinal fluid (CSF), white-matter (WM) and gray-matter (GM) was performed on the brain-extracted T1w using fast (FSL 5.0.9, RRID:SCR_002823, Zhang, Brady, and Smith 2001). Brain surfaces were reconstructed using recon-all (FreeSurfer 6.0.1, RRID:SCR_001847, (Dale et al., 1999)), and the brain mask estimated previously was refined with a custom variation of the method to reconcile ANTs-derived and FreeSurfer-derived

segmentations of the cortical gray-matter of Mindboggle (RRID:SCR_002438, (Klein et al., 2017)). Volume-based spatial normalization to one standard space (MNI152NLin2009cAsym) was performed through nonlinear registration with antsRegistration (ANTs 2.3.3), using brain-extracted versions of both T1w reference and the T1w template. The following template was selected for spatial normalization: ICBM 152 Nonlinear Asymmetrical template version 2009c [(Fonov et al., 2009), RRID:SCR_008796; TemplateFlow ID: MNI152NLin2009cAsym],

2.6.2.2 Functional data preprocessing

For each of the 5 BOLD runs found per subject (across all tasks and sessions), the following preprocessing was performed. First, a reference volume and its skull-stripped version were generated by aligning and averaging 1 single-band references (SBRefs). A deformation field to correct for susceptibility distortions was estimated based on fMRIPrep's fieldmap-less approach. The deformation field is that resulting from co-registering the BOLD reference to the same-subject T1w-reference with its intensity inverted (Huntenburg, 2014; Wang et al., 2017). Registration is performed with antsRegistration (ANTs 2.3.3), and the process regularized by constraining deformation to be nonzero only along the phase-encoding direction, and modulated with an average fieldmap template (Treiber et al., 2016). Based on the estimated susceptibility distortion, a corrected EPI (echo-planar imaging) reference was calculated for a more accurate co-registration with the anatomical reference. The BOLD reference was then co-registered to the T1w reference using bbregister (FreeSurfer) which implements boundary-based registration (Greve and Fischl, 2009). Co-registration was configured with six degrees of freedom. Head-motion parameters with respect to the BOLD reference (transformation matrices, and six corresponding rotation and translation parameters) are estimated before any spatiotemporal filtering using mcflirt (FSL 5.0.9, (Jenkinson et al., 2002)). BOLD runs were slice-time corrected using 3dTshift from AFNI 20160207 ((Cox and Hyde, 1997), RRID:SCR_005927). First, a reference volume and its skull-stripped version were generated using a custom methodology of fMRIPrep. The BOLD time-series (including slice-timing correction when applied) were resampled onto their original, native space by applying a single, composite transform to correct for head-motion and susceptibility distortions. These resampled BOLD time-series will be referred to as preprocessed BOLD in original space, or just preprocessed BOLD. The BOLD time-series were resampled into standard space, generating a preprocessed BOLD run in MNI152NLin2009cAsym space. First, a reference volume and its skull-stripped version were generated using a custom methodology of fMRIPrep. Several confounding time-series were calculated based on the preprocessed BOLD: framewise displacement (FD), DVARS and three region-wise global signals. FD was computed using two formulations following Power (absolute sum of relative motions, Power et al. (2014)) and

Jenkinson (relative root mean square displacement between affines, (Jenkinson et al., 2002)). FD and DVARS are calculated for each functional run, both using their implementations in Nipype (following the definitions by (Power et al., 2014)). The three global signals are extracted within the CSF, the WM, and the whole-brain masks. Additionally, a set of physiological regressors were extracted to allow for component-based noise correction (CompCor, (Behzadi et al., 2007)). Principal components are estimated after high-pass filtering the preprocessed BOLD time-series (using a discrete cosine filter with 128s cut-off) for the two CompCor variants: temporal (tCompCor) and anatomical (aCompCor). tCompCor components are then calculated from the top 2% variable voxels within the brain mask. For aCompCor, three probabilistic masks (CSF, WM and combined CSF+WM) are generated in anatomical space. The implementation differs from that of Behzadi et al. in that instead of eroding the masks by 2 pixels on BOLD space, the aCompCor masks are subtracted a mask of pixels that likely contain a volume fraction of GM. This mask is obtained by dilating a GM mask extracted from the FreeSurfer's aseg segmentation, and it ensures components are not extracted from voxels containing a minimal fraction of GM. Finally, these masks are resampled into BOLD space and binarized by thresholding at 0.99 (as in the original implementation). Components are also calculated separately within the WM and CSF masks. For each CompCor decomposition, the k components with the largest singular values are retained, such that the retained components' time series are sufficient to explain 50 percent of variance across the nuisance mask (CSF, WM, combined, or temporal). The remaining components are dropped from consideration. The head-motion estimates calculated in the correction step were also placed within the corresponding confounds file. The confound time series derived from head motion estimates and global signals were expanded with the inclusion of temporal derivatives and quadratic terms for each (Satterthwaite et al., 2013). Frames that exceeded a threshold of 0.5 mm FD or 1.5 standardized DVARS were annotated as motion outliers. All resampling can be performed with a single interpolation step by composing all the pertinent transformations (i.e., head-motion transform matrices, susceptibility distortion correction when available, and co-registrations to anatomical and output spaces). Gridded (volumetric) resamplings were performed using `antsApplyTransforms` (ANTs), configured with Lanczos interpolation to minimize the smoothing effects of other kernels (Lanczos, 1964). Non-gridded (surface) resamplings were performed using `mri_vol2surf` (FreeSurfer).

Many internal operations of fMRIPrep use Nilearn 0.6.2 ((Abraham et al., 2014), RRID:SCR_001362), mostly within the functional processing workflow. For more details of the pipeline, see the section corresponding to workflows in fMRIPrep's documentation¹².

2.6.2.3 General linear model

The fMRI data were analyzed using the general linear model (GLM) available in SPM12 – revision number 7771 (<https://www.fil.ion.ucl.ac.uk/spm/software/spm12/>) running on Matlab R2020b version. Ten regressors of interest (AccessMusicNo, AccessMusicYes, ElabMusicNo, ElabMusicYes, ChoiceMusic, AccessOdorNo, AccessOdorYes, ElabOdorNo, ElabOdorYes, ChoiceOdor) were defined based on the step of the memory process (*Access* when the participants perceived the cues, *Elab* when they were asked to remember their memories or to rest, and *Choice* when they selected a room and a period), on the sensory modality of the cue (*Music* or *Odor*) and on the Subjective Recollection response (“No” or “Yes”). The regressors were separately considered for each participant. They were independently modeled for each run by convolving condition-specific onset times with their duration (epochs) and a canonical hemodynamic response function (hrf). A high-pass filter eliminated instrumental and physiological low-frequency signal variations (the cutoff was determined as twice the maximum duration of a same event in a run), and temporal autocorrelation was modeled using an autoregressive process of order 1.

For model estimation, parameter estimates were obtained using the GLM in a voxelwise manner for each participant. The very preliminary functional analysis considered either the Access or the Elab phase. For Access, the contrast between the Yes and No Subjective recollection conditions [(AccessMusicYes – AccessMusicNo) and (AccessOdorYes – AccessOdorNo)] enabled the control over general sensory perception (sensory, attentional and motor aspects), visual instructions and the decision-making involved in subjective recollection response. For Elab, contrasting the Yes to No conditions [(ElabMusicYes – ElabMusicNo) and (ElabOdorYes – ElabOdorNo)] enabled the removal of brain activations arising from resting state. Significance was set at $p < 0.05$ FWE or $p < 0.001$ uncorrected otherwise. The human brain atlas of Mai et al. (2008) and the mean T1 image was used to localize and describe the activated regions. For presentation, the right side of the images correspond to the right side of the brain (neurological convention).

¹² Copyright Waiver: The above boilerplate text was automatically generated by fMRIPrep with the express intention that users should copy and paste this text into their manuscripts unchanged. It is released under the CC0 license.

2.6.3 Independent localizer (Odor detection / Music detection) tasks

Two independent fMRI runs, interleaved between the three runs dedicated to the main experiment, were used to localize odor-responsive and music-responsive cortical areas and to constrain functional regions of interest, as an unbiased method of identifying voxels for inclusion in the main functional analysis. In the odor localizer run, participants underwent an odor detection task consisting of 10 odor and 10 no-odor (air only) trials, using five odorants different from the main study, selected from (Rey et al., b, a) that were neutral, relatively unfamiliar: Basil (Créations aromatiques, Neuilly-Sur-Seine, France), Black Olive (Meilleur du Chef®, Bassussarry, France), Cis-3-hexenyl salicylate (Créations aromatiques), Dill (Pharmacie Croix Blanche, Dijon, France), Honey (Givaudan-Roure). Odors were presented the same way as in the main experiment. After 15 s of rest, odor and no-odor trials were presented every 15 s in pseudorandom order. Each trial began with the presentation of an instruction (“Odor?”). Participants pressed one of two buttons to indicate whether an odor was present or absent (left for “Yes”, right for “No”). Trials recurred with a stimulus-onset asynchrony of 15 s, and the run ended with a 15 s rest. A total of 220 functional volumes was acquired, using the same imaging parameters as for the main experiment. Image preprocessing and analysis were identical to the main experiment. After the model estimation in SPM12, odor *versus* no-odor conditions were contrasted, to generate statistical parametric maps of odor-responsive cortex, on a participant-by-participant basis. They will be used to delineate individual clusters which will be used later on to define an individual small volume of interests. Similar experimental design and statistical analyses were performed for the music detection localizer task. Musical excerpts were selected from the neutral stimulus from our previous experiment (Rey et al., a). They were neutral and relatively unfamiliar. The instruction beginning each trial was (“Music?”). Music *versus* no-music conditions were contrasted to generate music-responsive maps, as described for the odor stimuli.

3 Results

3.1 Behavioural results

A proportion of 8.8 ± 2.7 % of No responses was observed. The following calculations were performed with No responses removed.

3.1.1 Subjective recollection

The mean proportion of Yes response was of 73.3 ± 3.2 %. The 2 Sensory Modality (Music, Odor) x 2 Emotional Category (Negative, Positive) repeated measure ANOVA was performed on the proportion of subjectively remembered episode (Yes vs Yes + No responses) (Figure 2). It revealed a main effect of the cue’s Sensory Modality ($F_{(1,7)} = 9.22, p < .05$). The proportion of “Yes” was higher in the music (80.6 ± 5.2 %) than in the odor (66.0 ± 5.2 %) conditions. The

Emotional Category ($F_{(1,7)} = 6.04, p < .05$) also significantly influenced this proportion, the subjective recollection being higher in the positive ($79.2 \pm 5.2\%$) than in the negative ($67.4 \pm 5.2\%$) emotional conditions. The interaction was not significant ($F_{(1,7)} = 0.74, p = .79$).

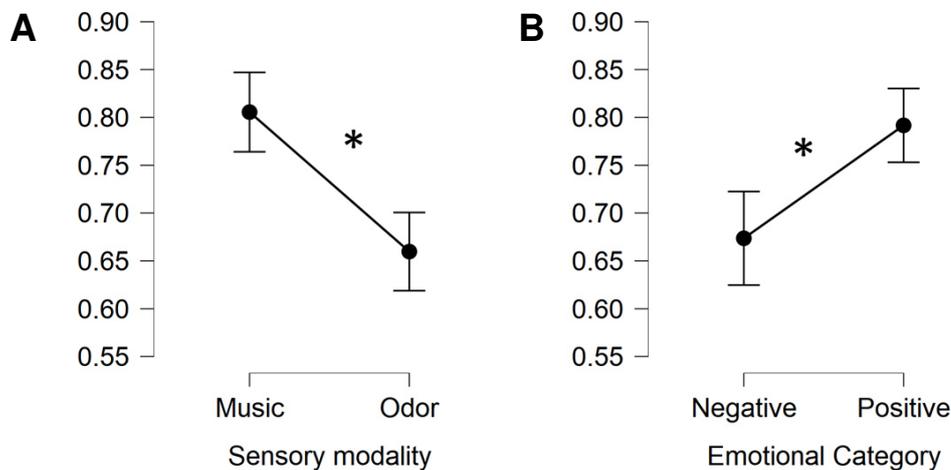


Figure 2. Subjective recollection. Mean proportion of Yes responses as a function of A. Sensory Modality, B. Emotional Category. *, $p < .05$.

3.1.2 Episodic Memory scores

First, a 2 Sensory Modality (Music, Odor) x 2 Emotional Category (Negative, Positive) x 4 Episodic Responses (Room-Period-, Room+Period-, Room-Period+, Room+Period+) repeated measure ANOVA was performed on the mean proportion of episodic responses. It revealed a significant main effect of the Episodic Responses ($F_{(3,7)} = 7.17, p_{Greenhouse-Geisser} < .05$), showing that the proportion of Room-Period- response ($39.2 \pm 3.4\%$) was higher than the proportion of the three other Episodic Responses (Room-Period+, $17.7 \pm 3.4\%$; Room+Period-, $17.9 \pm 3.4\%$; Room+Period+, $22.1 \pm 3.4\%$; $t_s > 3.17, p_s < .05$) (Figure 3). There were no other significant effects effect of the main factors or the interactions between factors ($F_s < 1.47, p_s > .26$).

Second, the proportion of Episodic Responses was compared to their respective chance level for each type of stimuli (negative music, positive music, negative odor, positive odor). Results demonstrated that only one proportion of responses was significantly above chance level: the accurate and complete memories evoked by positive odors (Room+Period+ for positive odor, $33.2 \pm 7.9\%$; $t_{(7)} = 2.82, p < .05$).

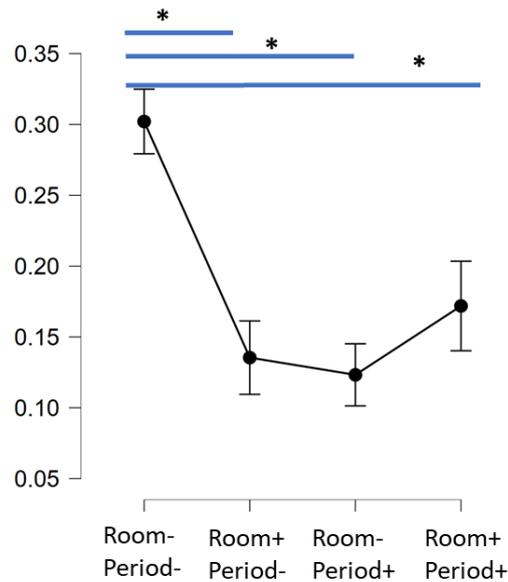


Figure 3. Episodic Responses. Mean proportion of responses as a function of Episodic Responses. *, $p < .05$.

3.1.3 Stimuli sensory evaluation

Each sensory evaluation was investigated with a 2 Sensory Modality (Music, Odor) x 2 Emotional Category (Negative, Positive) repeated measure ANOVA. Pleasantness significantly differed as a function of Emotional Category ($F_{(1,7)} = 32.54, p < .001$), positive stimuli (0.54 ± 0.09) being evaluated as being more pleasant than negative stimuli (-0.53 ± 0.09). This effect demonstrated the coherence of our selection, that is the *a priori* negative and positive stimuli. Emotional intensity was influenced by the the interaction between Emotional Category and Sensory Modality ($F_{(1,7)} = 6.18, p < .05$). Negative musical excerpts (0.46 ± 0.14) were evaluated as being more emotionally intense than negative odors ($-0.49 \pm 0.14, t = 4.79, p < .01$), while other contrasts were not significant (t 's $< 2.51, p$'s $> .23$). Familiarity tended to be significantly influenced by the Emotional Category ($F_{(1,7)} = 4.50, p = .072$), with a tendency for the positive stimuli (0.24 ± 0.12) to be evaluated as being more familiar than negative stimuli (-0.26 ± 0.12). Physical intensity was not significantly affected by the Sensory Modality and the Emotional Category. Describability was influenced by Emotional Category ($F_{(1,7)} = 6.50, p < .05$), positive stimuli (0.23 ± 0.09) being evaluated as being more describable than negative stimuli (-0.25 ± 0.09). Supplementary Figure 1 presented raw data stimuli sensory evaluation.

Functional results

Functional results presented here were only preliminary both in terms of group size (n=8) and in terms of the progress in functional data analysis. No formal conclusions should be made about the actual data presentation here.

First, the YesAccess condition was explored to check for data quality. Then, the influence of the feeling of recollection was investigated for each Sensory Modality and each Episodic Memory step (Access, Elab).

3.1.4 MusYAccess & OdoYAccess

Results are presented in Figure 4 (FWE $p < 0.05$). As expected, sensory brain regions were observed for both sensory modality conditions when participants were presented with the cues and declare they were able to access a memory. In Music condition, bilateral superior temporal gyrus was activated (left, size of the cluster in terms of the number of connected voxels (k) = 1807, x, y, z , Montreal Neurological Institute coordinates (in mm) of the maximum peak, -57 -15 4; right, k = 1995, 58 -15 7), corresponding the auditory cortex usually observed in music perception and memory (Janata, 2009a). In Odor condition, a small cluster in the left hemisphere going from the amygdala to the posterior piriform cortex (k = 66, -18, -4, -12) was observed, corresponding to primary olfactory brain regions observed in odor perception and memory (Saive et al., 2014b). Additionally, bilateral ant insula/post inferior frontal gyrus activation was observed in both Music (k = 648, -34 22 0; k = 332, 28 29 -3) and Odor (k = 101, -32 19 2; k = 97, 30 26 -3) conditions, suggesting access to the emotional memory. The inferior frontal gyrus is usually observed in both emotional processing (Lindquist et al., 2012) and autobiographical remembering (Greenberg et al., 2005b), and is suggested to be involved in both the enhancement and inhibition of emotional experience during autobiographical memory recollection (Denkova et al., 2013),

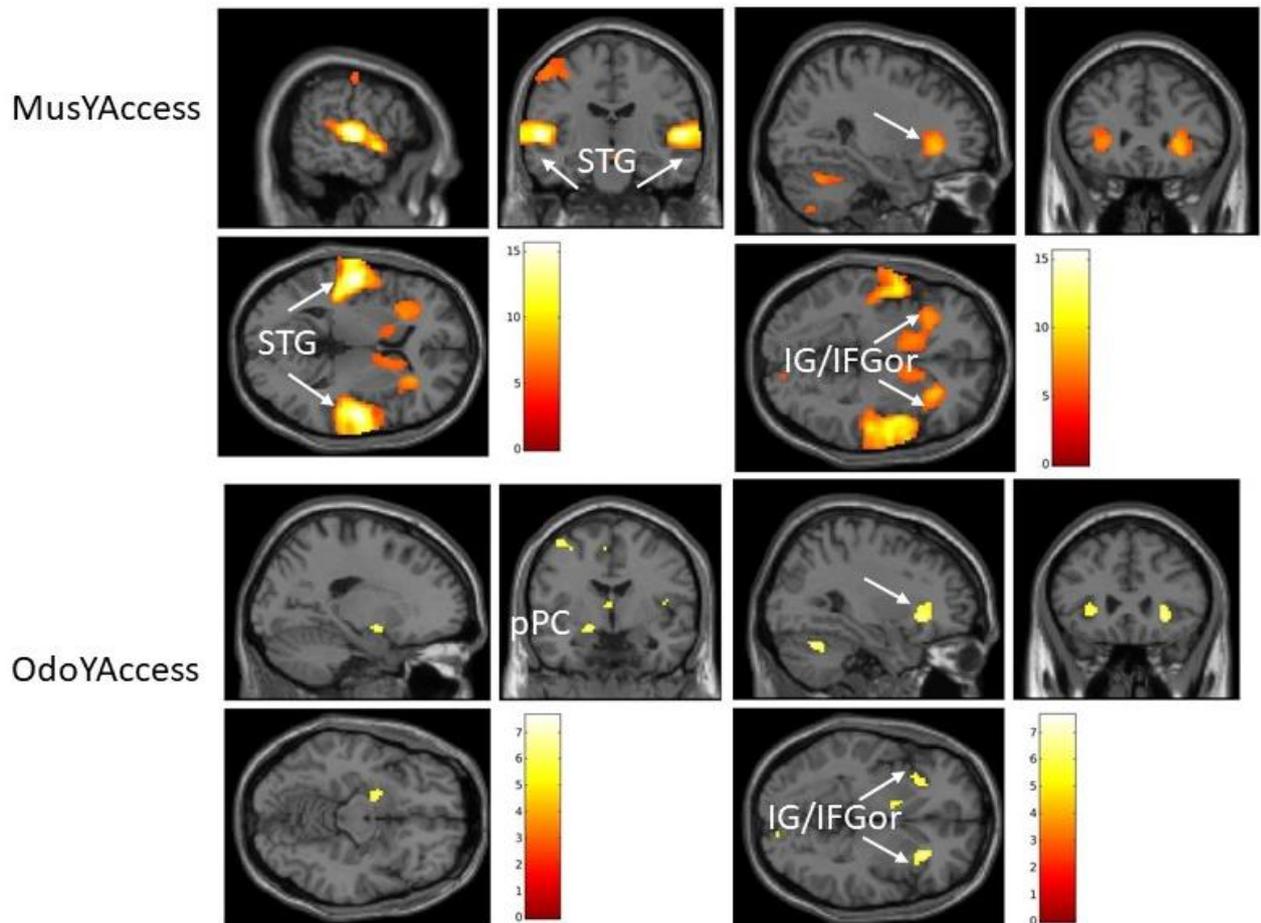


Figure 4. Brain areas engaged in the access of music-evoked (top) and odor-evoked (bottom) episodic memory. IG/IFGor, Insular gyrus / inferior frontal gyrus orbital part; pPC, posterior piriform cortex (extending to Amygdala); STG, superior temporal gyrus. The level of significance was set at $p < 0.05$ FWE corrected.

3.1.5 *MusYAccess-MusNAccess*

Access to a memory evoked by Music, in comparison with the feeling that no memory was accessed, was related to the activation of bilateral superior temporal gyrus (left, $k = 41$, x, y, z coordinates, $-57 -22 9$; right, $k = 56, 60 -13 7$) (Figure 5, FWE $p < 0.05$), a region that is consistently observed in music-evoked autobiographical memory retrieval (Janata, 2009a).

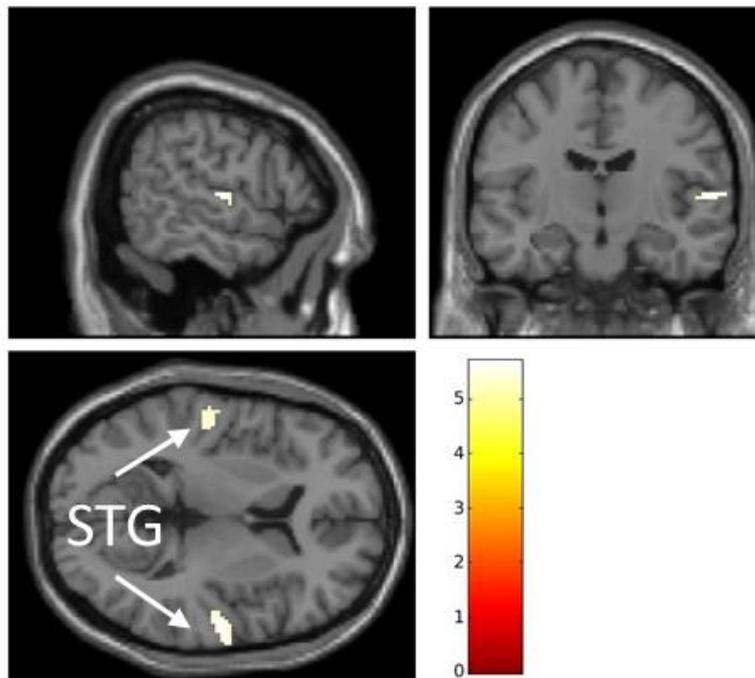


Figure 5. Brain areas engaged in the access of music-evoked episodic memory (Yes vs No subjective recollection). STG, superior temporal gyrus. The level of significance was set at $p < 0.05$ FWE corrected.

3.1.6 *OdoYAccess-OdoNAccess*

Access to a memory evoked by Odor, in comparison with the feeling that no memory was accessed, was not related to any significant activation, even at $p_{\text{uncorr}} < 0.001$.

3.1.7 *MusicYElab-MusicNElab*

Results are presented in Figure 6 (FWE $p < 0.05$). Elaboration of music-evoked memories, in comparison to rest, was related to left middle/superior frontal gyrus ($k = 47, -13, 6, 53$) activation. This brain regions is part of the medial prefrontal cortex, and is suggested to be a hub at which music, memories, and emotions are associated (Janata, 2009a). A left angular gyrus ($k = 42, -32 -61 43$) activation was also observed, this brain regions being central for the integration of features of the memories and therefore enabling the subjective experience of remembering (Bonnici et al., 2018).

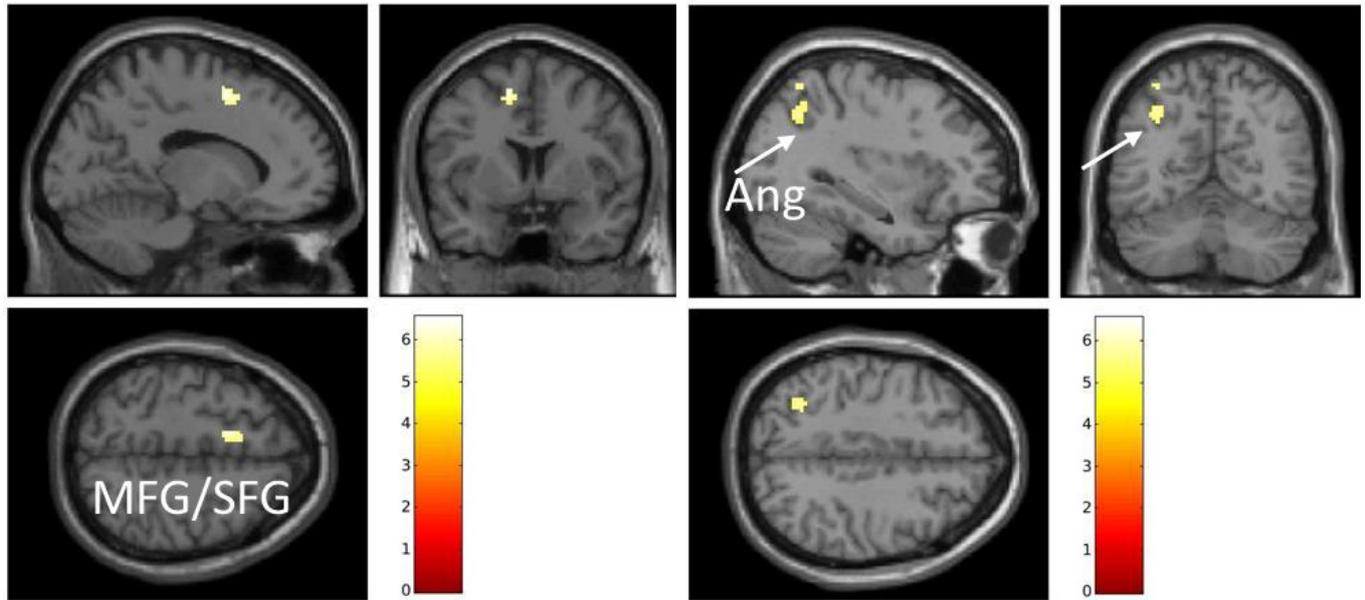


Figure 6. Brain areas engaged in the elaboration of music-evoked episodic memory (vs. Rest). Ang, angular gyrus; MFG/SFG, middle frontal gyrus / Superior frontal gyrus. The level of significance was set at $p < 0.05$ FWE corrected.

3.1.8 *OdorYElab-OdorNElab*

Results are presented in Figure 7 (FWE $p < 0.05$). Elaboration of odor-evoked memories, in comparison to rest, was related to similar brain regions than the ones observed in the elaboration of music-evoked memories, *i.e.*, the left middle frontal gyrus ($k = 205$, -37 1 62), left angular gyrus ($k = 41$, -32 -59 59). This is coherent with our previous work comparing neural substrates of odor and music memory (feeling of familiarity) where a high overlap was observed, including in these two brain regions (Plailly et al., 2007). A more posterior activation was also observed in the bilateral superior parietal lobule / angular gyrus (left, $k = 92$, -9 -73 57 ; right, $k = 50$, 14 -68 59) (Figure 7, FWE $p < 0.05$). It is a part of the dorsal parietal cortex, which is supposed to be involved in voluntary processes supporting retrieval search (Cabeza et al., 2008).

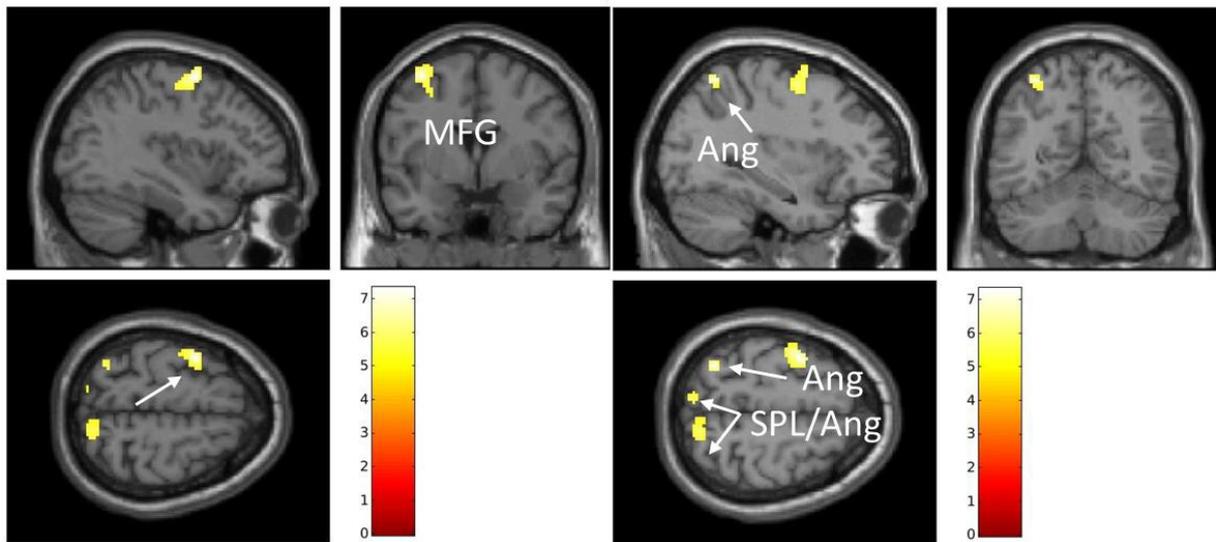


Figure 7. Brain areas engaged in the elaboration of odor-evoked episodic memory (vs. Rest). Ang, angular gyrus; MFG, middle frontal gyrus; SPL, superior parietal lobule. The level of significance was set at $p < 0.05$ FWE corrected.

4 Conclusions and perspectives

The aim of this study is to test the global hypothesis explaining that the singularity of odor-evoked memories comes from the close and direct anatomical links between the primary olfactory brain regions, memory and emotional structures (Daniels and Vermetten, 2016; Saive et al., 2014b). Particularly, the objective is to test whether the network that links the amygdala and the hippocampus with primary sensory areas is more strongly activated for the olfactory cue condition than the music cue condition, both cues being related to emotional and distinct memory of personal past events (Hackländer et al., 2018; Janata et al., 2007; Larsson et al., 2014). The study involves the encoding of complex events based on a What-Where-Which paradigm explored in virtual reality, and their retrieval in the fMRI scanner. The functional images acquired aimed to be analyzed with an effective connectivity technique, referred to as dynamic causal modeling. This technique will allow to study the effect of sensory modality and emotion on the network that links primary sensory areas identified with localizer runs, to the amygdala and the hippocampus.

Our results, although highly preliminary, were promising. The behavioral results indicate that the probability to subjectively recollect an episode was influenced by the emotion and the sensory modality of the cue. Music and positive stimuli were indeed followed by a higher feeling of recollection than odor and negative stimuli. However, most of yes responses involved, did not allow to recollect either the whole episode or at only one episodic dimension. These results thus suggest that music and positive stimuli may only be distinguishable from odor cue by the feeling of remembering they evoke. Whereas the emotion effect agree with research on this topic (Sharot et al., 2004), this hypothesis has never been tested with music stimuli.

Regarding functional data's, both access and elaboration phases of music-evoked EM and the elaboration phases of odor-evoked EM are followed by activation of regions involved in AM retrieval, here the left middle frontal gyrus and the left angular gyrus. The activation of these regions during the access for only music material may sustain the hypothesis that music is associated to the feeling of recollection early in the process, regardless of the memory accuracy. No activation was found for odor during the access, a result that may be linked with the poor ability of odor to directly access the feeling of remembering. At retrieval both regions involved in AM retrieval and in retrieval research (dorsal parietal cortex) for odors only, suggesting that the few cues that are able to be subjectively recollected induces an effective search process. Unfortunately, our data are at the moment insufficient to compare directly odor and music cues, and, more importantly the different network that are involved for the accurate and inaccurate access and elaboration phases.

5 Supplementary data

5.1 Sensory evaluations

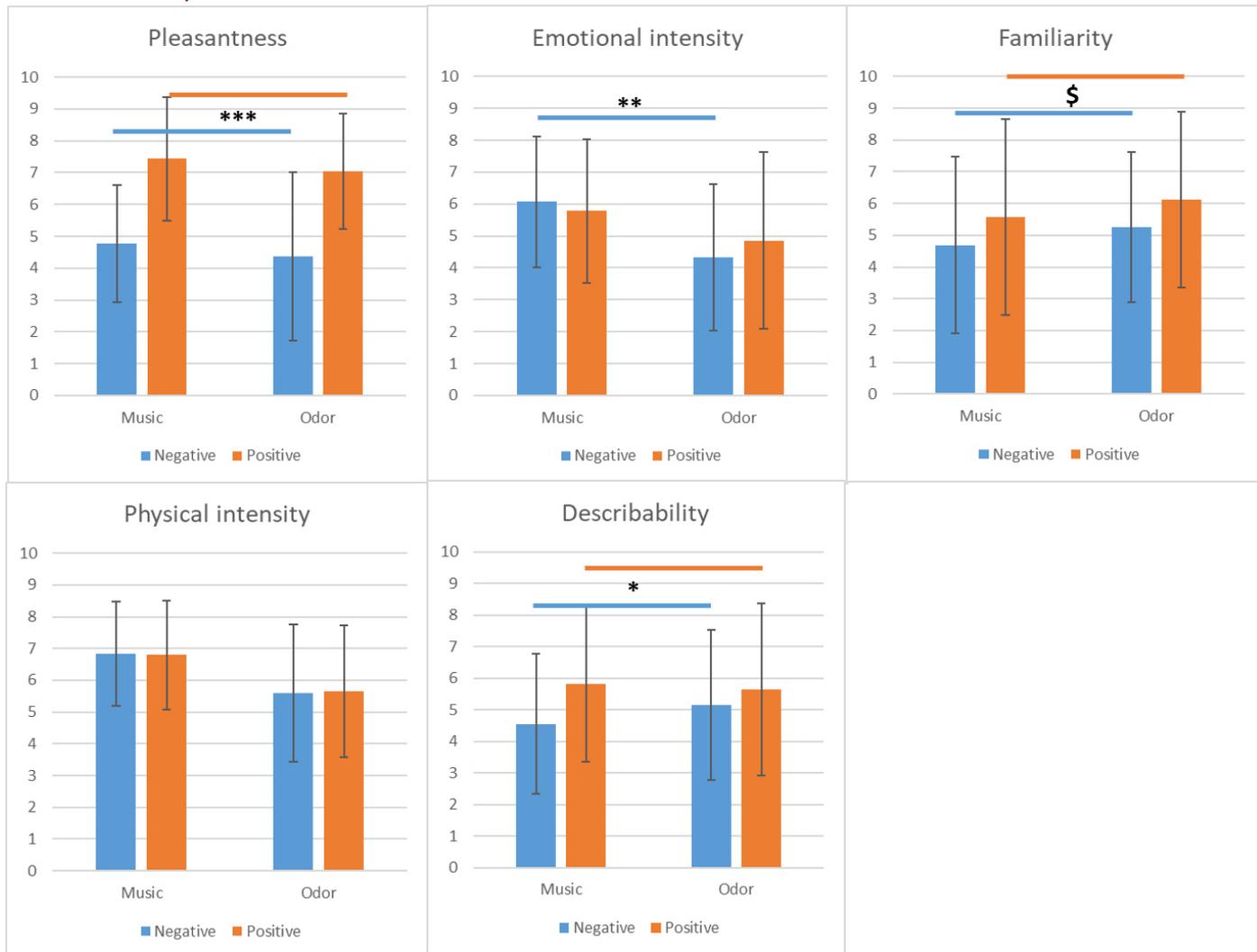


Figure 8. Sensory evaluations of the cues. Mean ratings as a function of Sensory Modality and Emotional Category. Error bars represent SD; *, $p < .001$; **, $p < .01$; *, $p < .05$; \$, $p < 0.08$.**

GENERAL DISCUSSION

Recollecting personal past events requires highly integrated and complex memory processes. The recall implies the reconstruction of the previously encoded events, which is highly sensitive to the environmental, physiological and psychological conditions under which the recall occurs. This reconstruction process brings to memory both its weakness, as the memory of past event is not the perfect copy of the encoding condition, and its force, as it confers to memory a highly adaptive value. Among the elements and dimensions that affect the retrieval process, the sensory modality of the recall-triggering cue influences both the nature of the retrieval and the content of the memory. Particularly, previous research has been shown that odor cues that evoke personal events in memory are singular in comparison to cues from other sensory modalities (Hackländer et al., 2018; Larsson et al., 2014), but the influence of the sensory modality of a cue remains still under investigated.

The general aim of this thesis was to unravel the sensory modality effect on memory of past events using a paradigm that allows the investigation of this memory in laboratory condition, which are closed to ecological ones. With the three studies of this thesis, we have shown that sensory modality is an important modulator of episodic retrieval. Further, we have demonstrated the differential effect of the emotion of the cue on episodic retrieval according to their sensory modality. Among the cue used, odors were the most powerful cue to trigger both episodic and associative memories, an ability that might be linked to their reward value. Music excerpts were highly recognized, an ability that allowed them to be associated to more autobiographical memories, but less being episodic. Faces were shown to be extremely good memory cues, an ability that may be linked to their ecological significance or their complexity.

We will begin by presenting the protocol we designed to study the recollection of personal past events in a controlled ecological way and by discussing what have been learned with its evolution during this thesis, as well as outlining its limits (1). Following this, we will provide a more complete view of the sensory modality effect revealed by our studies (2), and then discuss the differential influence of emotion on the sensory cue. Finally, we will overview the neural bases of these differences and discuss the potential perspectives and implications of the present research .

1 A new protocol to investigate episodic memory

To reveal the differential effect of the cue's sensory modality on episodic memory, different paradigms have been used in previous research. Historical episodic memory tasks, such as paradigms based on recall or recognition (e.g., the R/K paradigms Tulving, 1985) allow studying the memory quality of the encoded episode and eventually the associated subjective feelings. Autobiographical tasks allow for the complete and detailed study of the declarative memory content and associated subjective feelings (for example, see: Ernst *et al.*, 2021;

Jakubowski *et al.*, 2021; Levine *et al.*, 2002) . Even though these two sets of tasks offer a complete view of the involved episodic memory processes, the methodological and content-related gaps between them makes it difficult to integrate results from both approaches in a global conclusion. To bridge this gap, laboratory-ecological approaches have been started to set up (Pause *et al.*, 2013). Among them, the WWW paradigms appeared to be the most appropriate ones to study memory of personal events. These paradigms were constructed based on the Endel Tulving definition of past event memory (Clayton and Dickinson, 1998; Tulving, 1972). They allow for the control of the implicit encoding and the retrieval, while easily manipulating variables of interests. They are also associated to more subjective recollection than other episodic memory tasks, such as unexpected question/source memory or free recall tasks (Cheke and Clayton, 2013), and enable the *a posteriori* investigation of phenomenological characteristics of the displayed cues and/or events and/or the later investigation of these new, laboratory-created memories with the help of AM tasks. Finally, WWW tasks have allowed revealing that complete episodic retrieval is associated to high recollection performance. These findings have been observed in our laboratory in both non-human animals (Allerborn *et al.*, 2016; Veyrac *et al.*, 2015) and humans (Saive *et al.*, 2013, 2014a, 2015).

Based on the strength of the WWW tasks and our team's expertise, we have implemented the Episodor software in virtual reality environment. Two different protocols have been conceived based on the virtual exploration of a three-room house. In the first protocol version, Study 1, participants visited one room (bedroom, living room or office – *Which context*) per day. In each of the three rooms, three boxes that delivered an association of an odor, a music and a face (*What*) were placed at specific locations (*Where*). On the fourth day, during the retrieval task, participants were asked 1) whether they recognized the encoded cues among distractors, 2) whether they recalled the associated memory, and 3) to go, into the house, toward the room and select the box in which they thought the presented cue was displayed. This protocol allows us to compute the probabilities that the participants recollect the entire episode from the stimulus perception ($24.1 \pm 1.8 \%$) or from the stimulus recognition ($27.8 \pm 1.9 \%$). Results demonstrated their success in the task as they achieved it with a high accuracy given the associated chance level (respectively 5.5 % and 11 %), and the significant subjective feeling of recollection that was associated to the recall of accurate EM. It thus allows the investigation of the encoding and the retrieval of WWW events that lasted less than one day and that were associated with a subjective feeling of remembering, which is highly in agreement with the definition of episodic memory (Pause *et al.*, 2013). This first version of the protocol allowed us to investigate the sensory modality effect of the cue in Study 1 with complex, laboratory created memories, which are close to real-life events and their encoding. This first protocol

implementation nevertheless raised several questions calling for further improvement. In particular, one might wonder whether (1) the presentation of all stimuli together (*i.e.*, in a box) might impair the memory associated to each one separately, and (2) the implementation of the *where* and *which context* conditions were not too difficult and similar, notably with the boxes being located into the rooms that are themselves located into the house, leading to a two-level location task. This construction also involved that the participants had to remember the room to be able to select the right box. This features of the protocol prevented the investigation of the association memory between the cue and the *Where* dimension. In order to address these limitations, we designed a second version of Episodor in Study 2. Participants visited the entire house in three light conditions, one per day, defining the three contexts as the three periods of the day (nighttime, daytime, and twilight – *Which context*). In each room of the house (bedroom, living-room or office – *Where*) were located three objects, each delivering one stimulus of one sensory modality (*What*) when participants clicked on it. During the fourth day, participants were asked whether 1) they recognized the encoded cues among distractors, 2) they have the subjective feeling to recall the associated memory, and 3) to select, in the order they want, the picture of the room and the picture of the period of the day in which they thought they encountered the encoded cue. This paradigm also allows participants to recollect the entire episode. The computed probabilities, either from the stimulus perception ($11.4 \pm 1.7\%$) or the stimulus recognition ($14.6 \pm 2.0\%$) were quite high given the associated chance level (respectively 5.5% and 11%). The probabilities were also associated with an important subjective feeling of recollection. The discrepancies between the probabilities computed in Study 1 and Study 2 may be due to several reasons. First, in the protocol of Study 1 the participants were placed back into the virtual house to give their answer, whereas in the Study 2 the response was given by a click on the picture representing the conditions. In Study 1, participants were thus put back in the encoding context, a manipulation that may have helped them to recover the episode. This is similar to the previously reported context effect in word list recall (Godden and Baddeley, 1975; Smith *et al.*, 1978). However, one can also argue that this condition may have had no effect or only a weak effect as it was not a rich multi-sensory context but a predominantly visual context, (Wälti *et al.*, 2019). It is also possible that, in Study 2, the light context, together with the emotional context of the room that were changing each day, were insufficiently rich, to make the episode not identifiable enough. So, episodes may have been difficult to isolate from each other, leading to lower performance scores. Finally, as some participants reported an incongruity between the actual period of the day and the episode's period, the mis-match between real-world and the software condition may have make the encoding uneasy. Based on the observed findings and these potential influences, we could now implement a third version of the protocol to improve participant results while being even more ecological. For example, by running the same study and using the same

protocol as Study 2, but with other contexts such as the house style. Other solutions may be to choose a more immersive device (Smith, 2019) or to present the protocol in real condition may be another solutions.

Despite the difference in memory probabilities between the two protocols we used in Study 1 and Study 2, both versions rely strongly on episodic memory processes as defined by Endel Tulving (Tulving, 1972, 1985) and are ecologically valid, close to real-life event encoding?. However, the used virtual reality context still created a somewhat artificial situation. As laboratory and ecological tasks were shown to rely on different process (Cabeza *et al.*, 2004; Diamond *et al.*, 2020a; McDermott *et al.*, 2009; Monge *et al.*, 2018; Schnitzspahn *et al.*, 2011), future research could now further investigate whether and how the significance of our findings extend to even more ecological conditions (Foudil *et al.*, 2021). Our protocol implementation could be transposed in real life conditions, notably with participants being filmed when exploring a real house as implemented in a living lab setting (Dell'Era and Landoni, 2014), and thus we could further confirm its generalizability and ecological value. To bridge the gap between EM and AM studies, and to understand how they both allow for the investigation of personal event memory, it would be also interesting to ask participants to perform an AM test after the experiment. For example, it would be interesting to use the cueing method (for example, see Barzykowski *et al.*, 2019; Bonnici and Maguire, 2018), or to set up an interview method specific to this task. As AM tasks rely strongly on language whereas our task relies on motor responses, the study of personal event memory with both approaches may reveal information that allows a better understanding of the processes underlying memory of personal events. It may also allow to set up better strategies to help patients suffering from memory disorders such as Post-traumatic stress disorders or dementia.

2 Does the sensory modality of the memory cue influence episodic memory ?

2.1 Odors

The memory for odors showed approximately the same pattern of result in studies1 and 2. While being less frequently recognized than music and real faces, odors were able to evoke one or more dimensions of the to-be-recalled event, leading to the evocation of accurate episodic memories. Richard Davis (1977a) has previously shown similar results in an lab-based experiment in which odors were less recognized than abstract forms but were as potent as the forms to recall the associated digit. Thus, the step limiting the evocation of EM by odors is their recognition.

Odor recognition may be impaired relative to stimuli from other sensory modalities for several reasons. It is possible that the low attention that most humans pay to odors impairs their

recognition memory (McGann, 2017). It has indeed been shown that people who are more aware of odors in their everyday life are better able to recognize odors (Arshamian *et al.*, 2011). Odors are also poorly linked with the semantic memory system as demonstrated by the general poor ability to name odors and to talk about odors (Jraissati and Deroy, 2021; Majid and Burenhult, 2014). This poor link between odors and their potential name or associated concept labels may lead odors to be less recognized than other stimuli, as reflected in the finding that odors are better recognized when associated to a name than when difficult to name (Cessna and Frank, 2013; Cornell Karnekull *et al.*, 2015). The *level of processing theory* (Craik and Lockhart, 1972; Craik and Tulving, 1975) may explain why the weak link between odor perception and semanticization implies decreased recognition ability: the better the stimulus can be explained and categorized, the deeper will be its encoding and the better will be its recognition. The fact that odors are recognized less frequently than items of other sensory modalities may explain what is observed in AM studies, where odors were shown to evoke less memories than other types of cues (Ernst *et al.*, 2021; Goddard *et al.*, 2005; Hinton and Henley, 1993; Willander and Larsson, 2007). When less odors can be recognized, there is fewer chance for odors to cue a memory. The poorer number of memories cued by odors may also be associated to the mental blockage that the inability to name odors may induce. Participants may indeed be stuck on a research process of the odor associated label, preventing the access of episodic memory content (Hinton and Henley, 1993).

Despite the poor recognition ability associated to odors, odors are as powerful as visual stimuli to be associated to other objects (Davis, 1977; Herz, 1998; Herz and Cupchik, 1995; see Engen, 1987 for a review). This result was demonstrated by Study 1 and Study 2. Once an odor has been recognized, the chance to recover a memory that is specific and highly episodic is higher than those of music and virtual faces. These studies, together with other ran in a laboratory setting went further by showing that odors are able to evoke an entire, complex episode (as Saive *et al.*, 2013, 2014a, 2015, however with only odor stimuli), where other cues failed to evoke the episodic memories above chance level. These high episodic power of odor cues may be linked to some previous findings that have shown that odors evoke more specific and vivid memories (Larsson *et al.*, 2014). However, the cue sensory modality effect on AM is not often investigated by separating specific and episodic memories from other autobiographical memories. According to the Conway framework (Conway, 2005; Conway *et al.*, 2019), memories can be classified from the memories that are more semantized and global, to those that are more specific and episodic. For example, the memory of a three day festival in Clisson would be more global, generalized and constructed of global facts than the memory of the concert of Arch Enemy during this festival. So, when studies do not take the memory type into account (for example, see: Ernst *et al.*, 2021), they mix different types of

memories, which may hide potential effect. This may be particularly true for odor evoked memories, as they seem to evoke more specific memories

2.2 Music

In our studies, music-evoked memories present a different pattern of recognition and episodic responses than odor-evoked memories. The recognition of music pieces was very high, but the music pieces could hardly evoke associated dimensions and/or the full episodic memory. The recognition level for musical pieces was similar to that of real faces and higher than that of virtualized faces. This result is surprising, given that, for example, voices are less easily recognized compared to real faces (Damjanovic and Hanley, 2007). Concerning the music material, only three previous studies have compared the ability of music to be recognized to that of other sensory modality (visual and verbal stimuli; Cohen *et al.*, 2009, 2011; Deffler and Halpern, 2011). All three studies reported music to be less recognized than the other stimuli. Several reasons may explain the high recognition accuracy of music clips, the inconsistency with research showing a lower music recognition compared to visual cues. The relatively few number of cues (nine) that were used in all our experiments, may explain that both music and real faces are so well recognized that more stimuli are needed to make the difference (ceiling effect). This performance may have further been reinforced by the 5 mandatory presentations during encoding, an effect that have been shown to improve music recognition (Herff and Czernochowski, 2019). However, this could also be the case for odor and face cues and cannot thus be the only explanation.

In contrast to the high music recognition performance, our findings also demonstrated the weak ability of music to be associated to other episodic dimensions, coherently with research on music associative properties (Deffler and Halpern, 2011). However, other works have presented music as good AM cues. Music is indeed as good as odors or words (Herz, 1998; Janata *et al.*, 2007; Zator and Katz, 2017), or just a little bit weaker than face or television cues (Belfi *et al.*, 2016; Jakubowski *et al.*, 2021). A possibility may be that musical pieces are efficient to evoke AM because they are very well recognized. For example, music is still efficient to evoke complete episodic memories from their perception (from Cue to complete episode), in Study 1 and 2. Also, our studies 1 and 2 demonstrate the ability of music to induce a strong feeling of remembering that was comparable to other sensory modality in Study 2, and higher than for odors in Study 3. The Study 1 is an exception, with participant reporting more absence of subjective remembering than the presence of this feeling. This result may be explained, because music were selected to be neutral in Study 1, and emotion influences the feeling of remembering (Ferreri and Rodriguez-Fornells, 2017; Sharot *et al.*, 2004). Together, these results suggest that both the high potency of music to be recognized and the high feeling of remembering may further help memories to be recovered. Another explanation may be

provided for the weak potency of music to be an efficient episodic memory cue. Music might not have been sufficiently emotionally intense, and in particular rewarding. Indeed, previous research has shown that the reward value of a music piece improves the associated episodic recollection (Ferreri and Rodriguez-Fornells, 2017).

To conclude, music material seems to be a strong non-specific memory cue that is singularized by its strong ability to be recognized and to induce a strong feeling of recollection but leading to weak episodic memory recollection.

2.3 Faces

Faces were investigated in a real form and a virtual form in Studies 1 and 2, respectively. Faces were chosen because just like odors they are ecologically important stimulus for the human being (see the introduction chapter 3.1.1) and, in both real and virtual forms, they are processed holistically (Behrmann *et al.*, 2014; Maurer *et al.*, 2002; Taubert *et al.*, 2011), just like odors. Confirming the importance of faces for human being, our results demonstrate participants' high ability to recognize faces and to succeed in a complete episodic memory task, even with higher level than odors and music for Study 1. However, when faces are virtualized and thus treated as an out-group stimulus (*i.e.*, as other-ethnicity faces), they are recognized at the same level as odors and their evocation ability are considerably reduced in comparison to Study 1. These might be explained by the participants' strong expertise in face processing, even with unfamiliar faces, that may be shaped by their ecological importance. This expertise then leads faces to be a potent past event memory cue. Another possibility is that, if the a stimulus itself wears a lot of information, all these information's can be later used as different cues, or as cue reinforcer (Cessna and Frank, 2013; Cornell Karnekull *et al.*, 2015; Olsson *et al.*, 2009). The faces used in Study 1 were highly categorizable by differences in haircut, origin and gender for example. So, when they were presented as a cue, participant may have remembered the episode by remembering that an American-African female face with short hair was presented in this place in that context. It would thus not be the memory of the picture but the memory of the face's characteristics. Finally, the difference between the recognition of real and virtual faces may be that virtual faces were too simple to be able to be remembered at the same level as real faces. This may be surprising given that episodes were explored and retrieved – as in real life events – mostly visually (Greenberg *et al.*, 2005a), and given that even simple objects (*i.e.*, pictures of food, tools, animals) are able to evoke efficiently episodic memories (Foudil *et al.*, 2021; Silva *et al.*, 2020). However, contrary to simple objects, the here used faces were unfamiliar, unnamable, and when virtualized in Study 2, they had no ecological values and were less categorizable as all being created from male and female Caucasian faces with the same haircut and outfit. Thus, in Study 2 they followed a rather

shallow encoding processing, which might explain the reduced recognition scores (Craik and Lockhart, 1972; Craik and Tulving, 1975).

2.4 Potential explanations for differences observed between odor, music and face cues

The stimuli selected for these studies to investigate the comparison across the three sensory modalities, can all be distinguished by special features and phenomenological characteristics, and the way these stimuli are perceived shapes the characteristics of the memories they evoke.

1) Might the ecological relevance of the selected stimuli explain the difference between the odor, music and face cues ? Odor and real faces are ecologically relevant but this is less the case for music and virtual faces, which are mostly used as entertainment devices in everyday life (at least regarding the too simple music implemented here in this research). As, in Study 1, music and real faces recognition is similar, the ecological valence of stimuli may not influence the recognition process, or at a finer level than measured here. However, comparing the ability of cues to evoke an episodic memory, odors and real faces results are better than both music and virtual faces, suggesting that ecological relevance influences how personal events are recollected.

2) Might the ability to describe verbally a stimulus and its potential link with semantic processes improve both recognition and EM? According to the depth of processing hypothesis (Craik and Tulving, 1975), an easier way to describe and categorize stimulus may indeed lead to a better memory encoding and retrieval. In our research, for example, the change from Study 1 to Study 2, when faces changed from four to one face origin with the removal of defining features such as hairstyle, leads to the decrease of both recognition and EM performance. However, the ability to describe a stimulus seems to more strongly affect recognition than EM. To further investigate this potential influence, we ran a pilot study with one of our Master 1 student with 12 participants. These participants were told to write words that best described the stimulus. The number of words evoked by music and odor cues was compared (Figure 19). The results revealed that less words have been used to describe odors (4.83 ± 1.22 words) compared to music (6.53 ± 0.99 words), suggesting that odors were less describable than music. In Study 1 and Study 2, music was better recognized than odor but was less efficient as a memory cue, and as music was shown here to be more easily describable than odors, it is possible that the ability to describe an odor only influences recognition memory and not episodic memory. However, the memory of the 12 participants that took part to that pilot had not been investigated. To check this hypothesis, a new study needs to be run, in which participants would be required to describe the stimuli after the memory experiment.

Another possibility is that the better capacity to describe music can be involved in a successive cueing: a music being associated to one word, may evoke another word which is associated to an AM that is thus not directly related to the music cue. For example, a bouncing music played in piano may lead to the recognition of another bouncing music piece. Then, this second music piece will be able to induce the recall of the first moment this music piece was heard. As it is difficult to describe and categorize an odor, odors are less able to induce successive cueing. The additional fewer recognition performance and equal memory evocation feeling odors evoked compared to music cues, as we demonstrated in our studies, may lead to a lower ability to recall an AM relative to other senses. Odors seem thus to hold a special characteristic that mostly allows them to cue a complex episode. As odors have closed link with emotion, this may be the feature that distinguishes odor-evoked from music- and face-evoked EM.

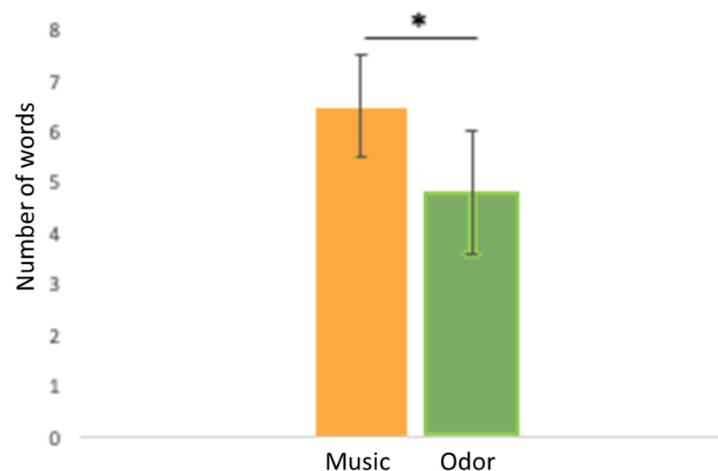


Figure 19: Sensory modality effect on the number of defining words

The presented results come from a pilot study ran by a M1 Student. The task was to describe the presented cue (odor in green and music in orange from the Study 3), with no other instructions and no time limits. An ANOVA (emotion category x Sensory modality) on the number of words revealed that music was defined with more words than odors without any emotional category effect.

3 Influence of sensory evaluations on memory performance as a function of sensory modality

Odor, music and faces are all emotional stimuli, in that they are able to potentially convey emotion. This dimension of those three types of stimuli may have considerable importance in that emotion may affect memory performance on many ways. For example, episodes are remembered differently when the emotion is traumatic (Hayes *et al.*, 2012), and highly emotionally intense stimuli favor the memory of the item compared to the memory of the context (Bisby *et al.*, 2016; Burke *et al.*, 1992; Mao *et al.*, 2015; Palombo *et al.*, 2018). This effect may be caused by an attentional catching effect of emotionally intense stimuli (Laney *et al.*, 2004), or because emotional arousal improves the feeling of remembering (in comparison to the feeling of knowing), but not the memory accuracy (Phelps and Sharot, 2008; Sharot *et*

al., 2004). However, most of these studies involved images that may be more arousing than images encountered in everyday-life situations, explaining that some studies using different types of stimuli did not replicate this effect (Cahill and McGaugh, 1995). AM, source and recollection studies have suggested that emotion has different effects on the access, the type and details of recalled personal events. According to Alisha Holland and Elizabeth Kensinger (Holland and Kensinger, 2010), the emotional intensity seems to improve the maintenance of detailed memories and the feeling of remembering while the valence would influence the access and the accuracy of a personal event memory. Moreover, the rewarding value was also shown to influence memory of personal memories (Ferreri and Rodriguez-Fornells, 2017; Ferreri *et al.*, 2021).

3.1 Is the emotion effect the same for each sensory modality?

In our studies, the influence of emotion was investigated by evaluating the effect of emotional intensity, valence and the wanting subjective ratings on recognition of the cue. The influence of emotion was investigated in all memory steps in Studies 1 to 3, but because of the currently still small number of participants in Study 3 we do not investigate emotion effect the same way as in Studies 1 and 2. So results for this last study were not reported here.

Emotional intensity was studied in Study 1 and Study 2. In those studies, both the probability to retrieve the entire episode from the stimulus perception and the probability to retrieve at least one episodic dimension were improved by the stimulus emotional intensity. In Study 1, designed with neutral stimuli, emotional intensity was shown to also influence the recognition process in an a-modal fashion, whereas in the Study 2 this influence only occurs for the musical material, and less for odor cues. In Studies 1 and 2, emotional intensity influenced only the recollection of the association between the recognized stimulus and one episodic dimension. This effect of emotional intensity may be linked to theories that propose an effect of familiarity on recollection by both its effect on the subjective feeling of recollection and the association to some, but not all, episodic dimensions (Phelps and Sharot, 2008; Sharot *et al.*, 2004). It thus would be interesting to study whether emotional intensity predict the probability to answer yes in the subjective recollection question, which is possible to do with the data collected from our three studies.

No emotional valence effect was shown in the first study, an expected result since the stimuli were selected to be neutral. In the second study; emotional valence influences in a u-shaped fashion the recognition and the recollection of the entire event from the perception and the recognition of the stimulus without any influence of the cue sensory modality. Both unpleasant and pleasant stimuli improved the recognition of the cues and the recollection of the entire event compared to the neutral stimuli. The valence effect was not shown to improve either the

memory of the context or the location. Thus, valence seems to influence the memory of the whole episode and not its elements.

The motivation to sample again a stimulus, here called the wanting, was only studied in the Study 2. The effect of wanting on recognition and episodic memory was different than the effect of pleasantness. It was surprising, given that negative stimuli were judged as being less pleasant wanted, the positive stimuli were more pleasant and wanted and the neutral stimuli were neither most nor less wanted and pleasant. The wanting effect on memory was modality dependent. For the recognition, the level of wanting influenced only the probability to recognized music targets: the more the music was wanted and the more it was recognized. Regarding episodic dimension, the wanting evaluations improved the probability to remember the association between the odor stimuli and all other episodic dimension, whether it was the period only, the room only or both. Wanting also tended to have an effect on face stimuli, improving the probability to remember the association between the period and the recognized face. No wanting effect was found for music stimuli and other faces episodic dimensions. As music and faces were only able to cue most of the episodic dimensions by chance, it was unlikely that an effect was found in this experiment. So, the result cannot rule out an eventual effect of wanting on these probabilities. However, as the wanting effect on episodic memory only appeared on odor stimuli, it is also possible to hypothesized that it is the wanting component of odor stimuli that allows them to be better associated to other episodic dimensions. Regarding the overall effect of the wanting dimension, it is interesting to note that the wanting, for odors and faces, only improves the associative memory from the recognition and not from the perception. Regarding the effect of wanting on music stimuli, it is thus possible that it is involved in two different processes: the recognition and the associative memory. This sustain the view the dual process theory (Yonelinas, 1994) presents: recognition may be done on two bases, either a familiarity process or a recollection process.

To conclude, emotions have a differential effect depending on sensory modality and the emotional scale used. Some emotion evaluations had a similar effect on episodic memory for all sensory modality, while other were dependent on the cue sensory modality. Valence and emotional intensity both influenced episodic memory and recognition regardless of the cue sensory modality. Valence influenced in a U-shaped fashion the recognition and the retrieval of the complete event, and may be linked to the recognition and recall accuracy. Emotional intensity may be related to the subjective feeling of remembering and promote the association between the stimulus and one episode dimension. The wanting influence seems to be modality dependent, or at least to more strongly affect odors than cues from other sensory modalities and to promote the association between the item and all episodic dimensions.

3.2 Is sensory modality influenced by other sensory dimensions ?

Previous research had shown that dimensions not related to emotion (*e.g.*, valence intensity, wanting, pleasantness, as implemented here) influence memory : the cue familiarity and complexity.

The cue familiarity influences recognition, associative memory and the subjective recollection feeling (Cornell Karnekull *et al.*, 2015; Davis, 1975; Saive *et al.*, 2015). In Study 1 and Study 2, familiarity was shown to influence music and real face recognition, and tended to do so for odors. Familiarity also improves the probability to recollect at least one episodic dimension, either from the stimulus perception or from the recognition for all sensory modalities. Our result confirmed the research on this topic (Cornell Karnekull *et al.*, 2015) and extend them by suggesting an effect of familiarity on associative memory. Together with the study of Anne-Lise Saive and collaborators (Saive *et al.*, 2015), taking into account that the evaluated stimuli of our three studies were all unfamiliar, and as the familiarity is influenced by the stimulus pleasantness (or valence Delplanque *et al.*, 2008), we can make the hypothesis that the familiarity effect may be mediated by the cue emotional intensity, pleasantness and/or valence.

The complexity evaluation of the cue was shown to influence recognition and episodic performance in Study 1, and confirmed research on this topic (Chai *et al.*, 2010; Craik and Tulving, 1975; Gilbert and Schleuder, 1990). In this study, it only improves the recognition of real faces and reduced the probability to form an association between the recognized stimulus and one episodic dimension whatever the sensory modality. The effect of complexity on face recognition may be related to the number of features wore by each faces, making them more recognizable among others. This is also possible that faces complexity allows encouraging a deeper encoding process, as demonstrated with words (Craik and Tulving, 1975). The deleterious effect of complexity on associative memory was surprising, as we expect that a more complex stimulus, associated to a higher processing depth, would be more recognized and associated to its context. However, the more the stimulus was complex and the less it was associated. One explanation may be that stimulus complexity increased the attention toward a stimulus and not the context, this favoring the memory of the items and reducing the memory of the context.

CONCLUSION

Episodic memory is a dynamic system and recall involves the reconstruction and recollection of previously encoded personal events. EM memory can thus be influenced by the recall environment. The research done in this thesis highlighted the influence of the cue's sensory modality on episodic memory retrieval in controlled conditions closed to the everyday-environment (**Study 1**). Resting on a *What-Where-Which* paradigms explored freely in a virtual-reality environment, we demonstrated that the sensory modality of the cue influenced episodic memory processes: similar cues from different sensory modalities are associated to different patterns of recognition and memory performance (**Studies 1 and 2**). Our results suggest that odors are the most efficient episodic memory cue. Although less well recognized than the other cues, they demonstrate high ability to cue all episodic memory dimensions. Music is highly recognized and poorly associated to episodic dimensions. Faces' data pattern changes whether they are real or virtualized: real faces show the best memory, as they demonstrated the higher recognition and episodic memory associated probabilities, whereas virtualized faces are averagely recognized and poorly evoke the EM dimensions.

We also demonstrated that stimuli from different sensory modalities are differently influenced by their phenomenological evaluations (**Study 2**). Importantly, EM was positively influenced by the rewarding value of the odor cue, whereas the rewarding value of the music cue only influenced music recognition.

The study of neural basis of EM cued by stimuli that different on their sensory modality and emotional category (**Study 3**) is currently still in progress and we present here some first preliminary analyses We notably shown that, only with 8 participants, regions significantly involved in AM retrieval are found in both access and retrieval of EM cued by both odors and faces.

This thesis may be the first step on a richer investigation of EM processes by the study of the sensory modality of the retrieval cue, as our studies raised numerous questions which some can be addressed in the perspective to this work.

Does the emotion effect come from the encoding or the retrieval? The research of this thesis focused on the sensory modality of the cue on episodic memory, that is during retrieval. However, the distinctive nature of EM evoked by odors may also be due to the effect at encoding, or through their higher ability to reinstate the encoding context. During the three experiments that were ran for this thesis, physiological measures had been taken at encoding (Study 3) and retrieval sessions (Study 1) and during the stimulus evaluation (Study 2). All these measures will us allow to understand how the stimulus phenomenological properties can contribute to the encoding and/or the recall of recognition and episodic information. For example, the analysis of physiological data taken when participants first explored the stimulus

at encoding, using the analytical tools we ran for Study 1 and 2, may reveal a different contribution of each evaluation on each retrieval probabilities, and it will be possible to compare the differential influence of the cue sensory modality at recall and at test to uncover their respective effects.

Investigation of the neural networks involved in EM cued by music and odors. The global hypothesis was that the power of odors to cue singular EM. This hypothesis comes from the close links between the primary olfactory sensory cortex, the amygdala and the hippocampus which have not been completely addressed in this thesis even though Study 3 is one step in this investigation. We are now planning to compare the networks engaged in odor-evoked EM and music-evoked EM, notably we aim to test this hypothesis with functional connectivity analyses, which will allow us to investigate the set of regions whose interactions are modulated by the sensory modality of the recall cue. Taking advantages of the differential effects of odor and music cues on EM, the study of the networks that are activated in the subjective recall and the elaboration phases in function of the sensory modality and the retrieval veracity may further reveal the specific networks that are engaged in the feeling of recollection without an accurate retrieval, and the feeling of recollection that is followed by an accurate retrieval.

Towards more ecological conditions of investigation. Although our paradigms take advantages of both a complex EM task and a virtual reality display, the conditions of investigation are not truly ecological. The use of simple stimuli in a task that takes place in the laboratory may have distorted the ecological validity of the results. It would thus be interesting to study more ecological material, in more ecological condition, for example by installing the experiment in a living lab (Dell’Era and Landoni, 2014). Another possibility is to investigate the participants’ AM with a cueing method of the encoded episodes either instead or after of the recall session. Recently, Alexandra Ernst, Herminie Bracq-Leca, Jane Plailly, Jean-Philippe Magué, and I ran a pilot for a new study with a Master student. We choose to compare specific AM evoked by meaningful odorants with pictures of objects of these odorant sources. The memory of participants is collected with a free verbalization task. It will be compared with analyses software allowing a rigorous and complete study of the participants’ discourse. Also, participants we be filmed and physiological responses such as heart rate and electrodermal activity will be recorded. This study, by its complexity and the wide variety of measures it will allow us to further complement the results obtained in this thesis.

Investigation of the effect of sensory modality in aging. With aging, the AM recall and the associated stimulus phenomenology change. For example, less memory is retrieved, with aging affecting more episodic than semantic memories. Aging leads to less episodic memory, the retrieval is less strongly based on sensory-perceptual features (Piolino et al., 2002, 2006;

Silva et al., 2020), and memory is differently affected by emotion (Narme et al., 2016). In the perceptual domain, odors and music hedonicity is changed, and odors are perceived as being less negative (Vieillard and Gilet, 2013; Vieillard et al., 2021). As most of our participants were young, our results are not generalizable to across the life span. Extending the study of the evolution of the sensory modality effect on the recognition and episodic memory pattern in an older population may lead to a better understanding of the evolution of memory. It also may be used to guide rehabilitation practices, giving indication on which type of cue can be used to address specific needs or re-evoked memories. For example, music can firstly be used to evoked many memories (Janata et al., 2007), and the supplementary use of odor cues may allow participant to evoke more dimensions.

REFERENCES

- Abichou, K., La Corte, V., Hubert, N., Orriols, E., Gaston-Bellegarde, A., Nicolas, S., and Piolino, P. (2019). Young and Older Adults Benefit From Sleep, but Not From Active Wakefulness for Memory Consolidation of What-Where-When Naturalistic Events. *Frontiers in Aging Neuroscience* 11, 58.
- Abraham, A., Pedregosa, F., Eickenberg, M., Gervais, P., Mueller, A., Kossaifi, J., Gramfort, A., Thirion, B., and Varoquaux, G. (2014). Machine learning for neuroimaging with scikit-learn. *Frontiers in Neuroinformatics* 8, 14.
- Allerborn, M., Gros, A., Messaoudi, B., Gervasoni, D., García, S., Thevenet, M., Laroche, S., Veyrac, A., and Ravel, N. (2016). A Novel Task for Studying Memory of Occasional Events in Rats. *Bio- Protocol* 6, e1740.
- Alonso, I., Dellacherie, D., and Samson, S. (2015). Emotional memory for musical excerpts in young and older adults. *Front Aging Neurosci* 7, 23.
- Altenmüller, E., Siggel, S., Mohammadi, B., Samii, A., and Münte, T.F. (2014). Play it again, Sam: brain correlates of emotional music recognition. *Front Psychol* 5.
- Andrade, J., and Donaldson, L. (2007). Evidence for an olfactory store in working memory? *An International Journal of Psychological Sciences* 50, 76–89.
- Angelucci, F.L., Silva, V.V., Pizzol, C.D., Spir, L.G., Praes, C.E.O., and Maibach, H. (2014). Physiological effect of olfactory stimuli inhalation in humans: an overview. *International Journal of Cosmetic Science* 36, 117–123.
- Annett, J.M. (1996). Olfactory Memory: A Case Study in Cognitive Psychology. *The Journal of Psychology* 130, 309–319.
- Armann, R.G.M., Jenkins, R., and Burton, A.M. (2016). A familiarity disadvantage for remembering specific images of faces. *J Exp Psychol Hum Percept Perform* 42, 571–580.
- Arnold, T.C., You, Y., Ding, M., Zuo, X.-N., Araujo, I. de, and Li, W. (2020). Functional Connectome Analyses Reveal the Human Olfactory Network Organization. *ENeuro* 7.
- Arshamian, A., and Larsson, M. (2014). Same same but different: the case of olfactory imagery. *Front. Psychol.* 0.
- Arshamian, A., Willander, J., and Larsson, M. (2011). Olfactory awareness is positively associated to odour memory. *Journal of Cognitive Psychology* 23, 220–226.
- Arshamian, A., Iannilli, E., Gerber, J.C., Willander, J., Persson, J., Seo, H.-S., Hummel, T., and Larsson, M. (2013). The functional neuroanatomy of odor evoked autobiographical memories cued by odors and words. *Neuropsychologia* 51, 123–131.
- Atkinson, R.C., and Shiffrin, R.M. (1968). Human Memory: A Proposed System and its Control Processes. In *Psychology of Learning and Motivation*, K.W. Spence, and J.T. Spence, eds. (Academic Press), pp. 89–195.
- Aubé, W., Peretz, I., and Armony, J.L. (2013). The effects of emotion on memory for music and vocalisations. *Memory* 21, 981–990.
- Avants, B.B., Epstein, C.L., Grossman, M., and Gee, J.C. (2008). Symmetric diffeomorphic image registration with cross-correlation: Evaluating automated labeling of elderly and neurodegenerative brain. *Medical Image Analysis* 12, 26–41.

- Baddeley, A. (2000). The episodic buffer: a new component of working memory? *Trends in Cognitive Sciences* 4, 417–423.
- Baddeley, A.D., and Hitch, G. (1974). Working Memory. In *Psychology of Learning and Motivation*, G.H. Bower, ed. (Academic Press), pp. 47–89.
- Baddeley, A.D., Hitch, G.J., and Allen, R.J. (2019). From short-term store to multicomponent working memory: The role of the modal model. *Mem Cogn* 47, 575–588.
- Balas, B., and Pacella, J. (2015). Artificial faces are harder to remember. *Comput Human Behav* 52, 331–337.
- Barrett, L.F., Adolphs, R., Marsella, S., Martinez, A.M., and Pollak, S.D. (2019). Emotional Expressions Reconsidered: Challenges to Inferring Emotion From Human Facial Movements: *Psychological Science in the Public Interest*.
- Bartolomei, F., Lagarde, S., Médina Villalon, S., McGonigal, A., and Benar, C.G. (2017). The “Proust phenomenon”: Odor-evoked autobiographical memories triggered by direct amygdala stimulation in human. *Cortex* 90, 173–175.
- Barzykowski, K., and Niedźwieńska, A. (2018). Priming involuntary autobiographical memories in the lab. *Memory* 26, 277–289.
- Barzykowski, K., and Staugaard, S.R. (2016). Does retrieval intentionality really matter? Similarities and differences between involuntary memories and directly and generatively retrieved voluntary memories. *British Journal of Psychology* 107, 519–536.
- Barzykowski, K., Niedźwieńska, A., and Mazzoni, G. (2019). How intention to retrieve a memory and expectation that a memory will come to mind influence the retrieval of autobiographical memories. *Consciousness and Cognition* 72, 31–48.
- Baudry, M. (2020). Did Proust predict the existence of episodic memory? *Neurobiology of Learning and Memory* 171, 107191.
- Becker, S., Bräscher, A.-K., Bannister, S., Bensafi, M., Calma-Birling, D., Chan, R.C.K., Eerola, T., Ellingsen, D.-M., Ferdenzi, C., Hanson, J.L., et al. (2019). The role of hedonics in the Human Affectome. *Neuroscience & Biobehavioral Reviews* 102, 221–241.
- Behrmann, M., Richler, J.J., Avidan, G., and Kimchi, R. (2014). Holistic face perception. In *Oxford Handbook of Perceptual Organization*, J. Wagemans, ed. (Oxford University Press), p.
- Behzadi, Y., Restom, K., Liu, J., and Liu, T.T. (2007). A component based noise correction method (CompCor) for BOLD and perfusion based fMRI. *NeuroImage* 37, 90–101.
- Belfi, A.M., Karlan, B., and Tranel, D. (2016). Music evokes vivid autobiographical memories. *Memory* 24, 979–989.
- Belfi, A.M., Karlan, B., and Tranel, D. (2018). Damage to the medial prefrontal cortex impairs music-evoked autobiographical memories. *Psychomusicology: Music, Mind, and Brain* 28, 201–208.
- Berntsen, D. (1996). Involuntary Autobiographical Memories. *Appl. Cognit. Psychol.* 10, 435–454.
- Berz, W.L. (1995). Working Memory in Music: A Theoretical Model. *Music Perception* 12, 353–364.

- Besson, G., Ceccaldi, M., and Barbeau, E.J. (2012). L'évaluation des processus de la mémoire de reconnaissance. *Revue de neuropsychologie Volume 4*, 242–254.
- Bisby, J.A., Horner, A.J., Hørlyck, L.D., and Burgess, N. (2016). Opposing effects of negative emotion on amygdalar and hippocampal memory for items and associations. *Social Cognitive and Affective Neuroscience 11*, 981–990.
- Blair, R.J.R. (2003). Facial expressions, their communicatory functions and neuro-cognitive substrates. *Philos Trans R Soc Lond B Biol Sci 358*, 561–572.
- Bonnici, H.M., and Maguire, E.A. (2018). Two years later – Revisiting autobiographical memory representations in vmPFC and hippocampus. *Neuropsychologia 110*, 159–169.
- Bonnici, H.M., Cheke, L.G., Green, D.A.E., FitzGerald, T.H.M.B., and Simons, J.S. (2018). Specifying a Causal Role for Angular Gyrus in Autobiographical Memory. *J Neurosci 38*, 10438–10443.
- Boyle, J.A., Lundström, J.N., Knecht, M., Jones-Gotman, M., Schaal, B., and Hummel, T. (2006). On the trigeminal percept of androstenone and its implications on the rate of specific anosmia. *Journal of Neurobiology 66*, 1501–1510.
- Brédart, S., and Barsics, C. (2012a). Recalling Semantic and Episodic Information From Faces and Voices: A Face Advantage. *Current Directions in Psychological Science 21*, 378–381.
- Brédart, S., and Barsics, C. (2012b). Recalling Semantic and Episodic Information From Faces and Voices: A Face Advantage. *Current Directions in Psychological Science 21*, 378–381.
- Bruce, V. (1982). Changing faces: Visual and non-visual coding processes in face recognition. *British Journal of Psychology 73*, 105–116.
- de Bruijn, M.J., and Bender, M. (2017). Olfactory cues are more effective than visual cues in experimentally triggering autobiographical memories. *Memory 1–12*.
- Burke, A., Heuer, F., and Reisberg, D. (1992). Remembering emotional events. *Memory & Cognition 20*, 277–290.
- Burton, A.M., White, D., and McNeill, A. (2010). The Glasgow Face Matching Test. *Behavior Research Methods 42*, 286–291.
- Bushdid, C., Magnasco, M.O., Vosshall, L.B., and Keller, A. (2014). Humans Can Discriminate More than 1 Trillion Olfactory Stimuli. *Science 343*, 1370–1372.
- Cabanac, M. (1971). Physiological Role of Pleasure. *Science 173*, 1103–1107.
- Cabeza, R., and Moscovitch, M. (2013). Memory Systems, Processing Modes, and Components: Functional Neuroimaging Evidence. *Perspect Psychol Sci 8*, 49–55.
- Cabeza, R., and St Jacques, P. (2007). Functional neuroimaging of autobiographical memory. *Trends Cogn Sci 11*, 219–227.
- Cabeza, R., Prince, S., Daselaar, S., Greenberg, D., Budde, M., Dolcos, F., LaBar, K., and Rubin, D. (2004). Brain Activity during Episodic Retrieval of Autobiographical and Laboratory Events: An fMRI Study using a Novel Photo Paradigm. *Journal of Cognitive Neuroscience 16*, 1583–1594.

- Cabeza, R., Ciaramelli, E., Olson, I.R., and Moscovitch, M. (2008). Parietal Cortex and Episodic Memory: An Attentional Account. *Nat Rev Neurosci* 9, 613–625.
- Cady, E.T., Harris, R.J., and Knappenberger, J.B. (2008). Using music to cue autobiographical memories of different lifetime periods. *Psychology of Music* 36, 157–177.
- Cahill, L., and McGaugh, J.L. (1995). A Novel Demonstration of Enhanced Memory Associated with Emotional Arousal. *Consciousness and Cognition* 4, 410–421.
- Cahill, L., Babinsky, R., Markowitsch, H.J., and McGaugh, J.L. (1995). The amygdala and emotional memory. *Nature* 377, 295–296.
- Cameron, E.L., Møller, P., and Karn, K.S. (2021). Effects of COVID-19 on Sense of Smell: Human Factors/Ergonomics Considerations. *Hum Factors* 0018720821990162.
- Camina, E., and Güell, F. (2017). The Neuroanatomical, Neurophysiological and Psychological Basis of Memory: Current Models and Their Origins. *Front. Pharmacol.* 8.
- Carbon, C.-C., and Leder, H. (2005). When Feature Information Comes First! Early Processing of Inverted Faces. *Perception* 34, 1117–1134.
- Cessna, T., and Frank, R. (2013). Does Odor Knowledge or an Odor Naming Strategy Mediate the Relationship Between Odor Naming and Recognition Memory? *Chemosensory Perception*.
- Chai, X., Ofen, N., Jacobs, L., and Gabrieli, J. (2010). Scene complexity: influence on perception, memory, and development in the medial temporal lobe. *Frontiers in Human Neuroscience* 4, 21.
- Cheke, L.G., and Clayton, N.S. (2013). Do different tests of episodic memory produce consistent results in human adults? *Learn. Mem.* 20, 491–498.
- Chen, H.-Y., Gilmore, A.W., Nelson, S.M., and McDermott, K.B. (2017). Are There Multiple Kinds of Episodic Memory? An fMRI Investigation Comparing Autobiographical and Recognition Memory Tasks. *The Journal of Neuroscience* 37, 2764–2775.
- Chen, W., Liu, C.H., Li, H., Tong, K., Ren, N., and Fu, X. (2015). Facial expression at retrieval affects recognition of facial identity. *Frontiers in Psychology* 6, 780.
- Cheung, V.K.M., Harrison, P.M.C., Meyer, L., Pearce, M.T., Haynes, J.-D., and Koelsch, S. (2019). Uncertainty and Surprise Jointly Predict Musical Pleasure and Amygdala, Hippocampus, and Auditory Cortex Activity. *Current Biology* 29, 4084-4092.e4.
- Chu, S., and Downes, J. (2000a). Long live Proust: The odour-cued autobiographical memory bump. *Cognition* 75, B41–B50.
- Chu, S., and Downes, J.J. (2000b). Odour-evoked Autobiographical Memories: Psychological Investigations of Proustian Phenomena. *Chem Senses* 25, 111–116.
- Chu, S., and Downes, J.J. (2002). Proust nose best: Odors are better cues of autobiographical memory. *Memory & Cognition* 30, 511–518.
- Clayton, N.S., and Dickinson, A. (1998). Episodic-like memory during cache recovery by scrub jays. *Nature* 395, 272–274.

- Cohen, M., Horowitz, T., and Wolfe, J. (2009). Auditory recognition memory is inferior to visual recognition memory. *Proceedings of the National Academy of Sciences of the United States of America* *106*, 6008–6010.
- Cohen, M.A., Evans, K.K., Horowitz, T.S., and Wolfe, J.M. (2011). Auditory and visual memory in musicians and nonmusicians. *Psychon Bull Rev* *18*, 586–591.
- Conway, M.A. (2005). Memory and the self. *Journal of Memory and Language* *53*, 594–628.
- Conway, M.A. (2009). Episodic memories. *Neuropsychologia* *47*, 2305–2313.
- Conway, M.A., and Pleydell-Pearce, C.W. (2000). The construction of autobiographical memories in the self-memory system. *Psychological Review* *107*, 261–288.
- Conway, M.A., Justice, L.V., and D'Argembeau, A. (2019). The self-memory system revisited: Past, present, and future. In *The Organization and Structure of Autobiographical Memory*, (New York, NY, US: Oxford University Press), pp. 28–51.
- Cornell Karnekull, S., Jonsson, F.U., Willander, J., Sikstrom, S., and Larsson, M. (2015). Long-Term Memory for Odors: Influences of Familiarity and Identification Across 64 Days. *Chemical Senses* *40*, 259–267.
- Cornell Karnekull, S., Arshamian, A., Nilsson, M.E., and Larsson, M. (2018). The Effect of Blindness on Long-Term Episodic Memory for Odors and Sounds. *Front. Psychol.* *9*.
- Corrow, S.L., Dalrymple, K.A., and Barton, J.J. (2016). Prosopagnosia: current perspectives. *EB* *8*, 165–175.
- Cortes, D.S., Laukka, P., Lindahl, C., and Fischer, H. (2017). Memory for faces and voices varies as a function of sex and expressed emotion. *PLoS ONE* *12*, e0178423.
- Cox, R.W., and Hyde, J.S. (1997). Software tools for analysis and visualization of fMRI data. *NMR in Biomedicine* *10*, 171–178.
- Craik, F.I.M. (2020). Remembering: An Activity of Mind and Brain. *Annu. Rev. Psychol.* *71*, 1–24.
- Craik, F.I.M., and Lockhart, R.S. (1972). Levels of processing: A framework for memory research. *Journal of Verbal Learning and Verbal Behavior* *11*, 671–684.
- Craik, F.I.M., and Tulving, E. (1975). Depth of processing and the retention of words in episodic memory. *Journal of Experimental Psychology: General* *104*, 268–294.
- Cross, I. (2012). Cognitive Science and the Cultural Nature of Music. *Topics in Cognitive Science* *4*, 668–677.
- Cross, I. (2014). Music and communication in music psychology. *Psychology of Music* *42*, 809–819.
- Crouzet, S.M., Kirchner, H., and Thorpe, S.J. (2010). Fast saccades toward faces: Face detection in just 100 ms. *Journal of Vision* *10*, 16.
- Croy, I., Nordin, S., and Hummel, T. (2014). Olfactory Disorders and Quality of Life—An Updated Review. *Chemical Senses* *39*, 185–194.

- Cuddy, L.L. (2018). Long-Term Memory for Music. In Springer Handbook of Systematic Musicology, R. Bader, ed. (Berlin, Heidelberg: Springer), pp. 453–459.
- Dale, A.M., Fischl, B., and Sereno, M.I. (1999). Cortical Surface-Based Analysis: I. Segmentation and Surface Reconstruction. *NeuroImage* 9, 179–194.
- Damjanovic, L., and Hanley, J.R. (2007). Recalling episodic and semantic information about famous faces and voices. *Memory & Cognition* 35, 1205–1210.
- Daniels, J.K., and Vermetten, E. (2016). Odor-induced recall of emotional memories in PTSD—Review and new paradigm for research. *Experimental Neurology* 284, 168–180.
- D’Argembeau, A., and Van der Linden, M. (2007). Facial expressions of emotion influence memory for facial identity in an automatic way. *Emotion* 7, 507–515.
- Davis, R.G. (1975). Acquisition of verbal associations to olfactory stimuli of varying familiarity and to abstract visual stimuli. *Journal of Experimental Psychology: Human Learning and Memory* 1, 134–142.
- Davis, R.G. (1977a). Acquisition and retention of verbal associations to olfactory and abstract visual stimuli of varying similarity. *Journal of Experimental Psychology: Human Learning and Memory* 3, 37–51.
- Davis, R.G. (1977b). Acquisition and retention of verbal associations to olfactory and abstract visual stimuli of varying similarity. *J. Exp. Psychol.-Hum. L.* 3, 37–51.
- Deffler, S.A., and Halpern, A.R. (2011). Contextual information and memory for unfamiliar tunes in older and younger adults. *Psychology and Aging* 26, 900–904.
- Dell’Era, C., and Landoni, P. (2014). Living Lab: A Methodology between User-Centred Design and Participatory Design. *Creativity and Innovation Management* 23, 137–154.
- Delplanque, S., Grandjean, D., Chrea, C., Aymard, L., Cayeux, I., Le Calvé, B., Velazco, M.I., Scherer, K.R., and Sander, D. (2008). Emotional Processing of Odors: Evidence for a Nonlinear Relation between Pleasantness and Familiarity Evaluations. *Chemical Senses* 33, 469–479.
- Demorest, S.M., Morrison, S.J., Stambaugh, L.A., Beken, M., Richards, T.L., and Johnson, C. (2010). An fMRI investigation of the cultural specificity of music memory. *Soc. Cogn. Affect. Neurosci.* 5, 282–291.
- Denkova, E., Dolcos, S., and Dolcos, F. (2013). The Effect of Retrieval Focus and Emotional Valence on the Inferior Frontal Cortex Activity during Autobiographical Recollection. *Frontiers in Behavioral Neuroscience* 7, 192.
- Diamond, N.B., Abdi, H., and Levine, B. (2020a). Different patterns of recollection for matched real-world and laboratory-based episodes in younger and older adults. *Cognition* 202, 104309.
- Diamond, N.B., Armson, M.J., and Levine, B. (2020b). The Truth Is Out There: Accuracy in Recall of Verifiable Real-World Events. *Psychol Sci* 31, 1544–1556.
- Dickerson, B.C., and Eichenbaum, H. (2010). The Episodic Memory System: Neurocircuitry and Disorders. *Neuropsychopharmacol* 35, 86–104.
- Dubourg, L., Silva, A.R., Fitamen, C., Moulin, C.J.A., and Souchay, C. (2016). SenseCam: A new tool for memory rehabilitation? *Revue Neurologique* 172, 735–747.

- Duchaine, B., and Yovel, G. (2015). A Revised Neural Framework for Face Processing. *Annual Review of Vision Science* 1, 393–416.
- Dudai, Y. (2004). The Neurobiology of Consolidations, Or, How Stable is the Engram? *Annu. Rev. Psychol.* 55, 51–86.
- Dudai, Y., Karni, A., and Born, J. (2015). The Consolidation and Transformation of Memory. *Neuron* 88, 20–32.
- Dyck, M., Winbeck, M., Leiberg, S., Chen, Y., Gur, R.C., and Mathiak, K. (2008). Recognition Profile of Emotions in Natural and Virtual Faces. *PLoS ONE* 3, e3628.
- Eacott, M.J., and Easton, A. (2010). Episodic memory in animals: Remembering which occasion. *Neuropsychologia* 48, 2273–2280.
- Easton, A., and Eacott, M.J. (2008). Chapter 2.3 A new working definition of episodic memory: replacing “when” with “which.” In *Handbook of Behavioral Neuroscience*, (Elsevier), pp. 185–196.
- Easton, A., Webster, L.A.D., and Eacott, M.J. (2012). The episodic nature of episodic-like memories. *Learning & Memory* 19, 146–150.
- Ebbinghaus, H. (1885). *Über das Gedächtnis* (Leipzig: Duncker & Humblot).
- Ebbinghaus, H., Ruger, H.A., and Bussenius, C.E. (1913). *Memory; a contribution to experimental psychology* (New York City: Teachers College, Columbia University).
- Eichenbaum, H., Yonelinas, A. p., and Ranganath, C. (2007). The Medial Temporal Lobe and Recognition Memory. *Annu. Rev. Neurosci.* 30, 123–152.
- Ekman, P., and Friesen, W.V. (1976). Measuring facial movement. *J Nonverbal Behav* 1, 56–75.
- El Haj, M., Fasotti, L., and Allain, P. (2012). The involuntary nature of music-evoked autobiographical memories in Alzheimer’s disease. *Consciousness and Cognition* 21, 238–246.
- EL Haj, M., Gandolphe, M.C., Gallouj, K., Kapogiannis, D., and Antoine, P. (2018). From Nose to Memory: The Involuntary Nature of Odor-evoked Autobiographical Memories in Alzheimer’s Disease. *Chem Senses* 43, 27–34.
- Engen, T. (1987). Remembering Odors and Their Names. *American Scientist* 75, 497–503.
- Ernst, A., Bertrand, J.M.F., Voltzenlogel, V., Souchay, C., and Moulin, C.J.A. (2021). The Proust Machine: What a Public Science Event Tells Us About Autobiographical Memory and the Five Senses. *Front. Psychol.* 11.
- Eschrich, S. (2005). Remember Bach: An Investigation in Episodic Memory for Music. *Annals of the New York Academy of Sciences* 1060, 438–442.
- Eschrich, S., Münte, T.F., and Altenmüller, E.O. (2008). Unforgettable film music: The role of emotion in episodic long-term memory for music. *BMC Neuroscience* 9, 48.
- Esfahani-Bayerl, N., Finke, C., Kopp, U., Moon, D.-U., and Ploner, C.J. (2019). Musical memory and hippocampus revisited: Evidence from a musical layperson with highly selective hippocampal damage. *Cortex* 119, 519–527.

- Esteban, O., Markiewicz, C.J., Goncalves, M., DuPre, E., Kent, J.D., Salo, T., Ciric, R., Pinsard, B., Blair, R.W., Poldrack, R.A., et al. (2021a). fMRIprep: a robust preprocessing pipeline for functional MRI (Zenodo).
- Esteban, O., Markiewicz, C.J., Burns, C., Goncalves, M., Jarecka, D., Ziegler, E., Berleant, S., Ellis, D.G., Pinsard, B., Madison, C., et al. (2021b). nipy/nipype: 1.7.0 (Zenodo).
- Eustache, F., and Desgranges, B. (2008). MNESIS: towards the integration of current multisystem models of memory. *Neuropsychol Rev* 18, 53–69.
- Eustache, F., and Desgranges, B. (2010). Les chemins de la mémoire.
- Eustache, F., Viard, A., and Desgranges, B. (2016). The MNESIS model: Memory systems and processes, identity and future thinking. *Neuropsychologia* 87, 96–109.
- Feingold, G.A. (1914). The Influence of Environment on Identification of Persons and Things.. *Journal of the American Institute of Criminal Law & Criminology* 5, 39–51.
- Ferbinteanu, J. (2019). Memory systems 2018 – Towards a new paradigm. *Neurobiology of Learning and Memory* 157, 61–78.
- Ferdenzi, C., Richard Ortégón, S., Delplanque, S., Baldovini, N., and Bensafi, M. (2020). Interdisciplinary challenges for elucidating human olfactory attractiveness. *Philosophical Transactions of the Royal Society B: Biological Sciences* 375, 20190268.
- Ferreri, L., and Rodriguez-Fornells, A. (2017). Music-related reward responses predict episodic memory performance. *Experimental Brain Research*.
- Ferreri, L., Mas-Herrero, E., Cardona, G., Zatorre, R.J., Antonijoan, R.M., Valle, M., Riba, J., Ripollés, P., and Rodriguez-Fornells, A. (2021). Dopamine modulations of reward-driven music memory consolidation. *Ann N Y Acad Sci*.
- Fitch, W.T., and Popescu, T. (2019). The world in a song. *Science*.
- Fiveash, A., Bedoin, N., Gordon, R.L., and Tillmann, B. (2021). Processing rhythm in speech and music: Shared mechanisms and implications for developmental speech and language disorders. *Neuropsychology*.
- Fonov, V., Evans, A., McKinstry, R., Almlí, C., and Collins, D. (2009). Unbiased nonlinear average age-appropriate brain templates from birth to adulthood. *NeuroImage* 47, S102.
- Ford, J.H., Addis, D.R., and Giovanello, K.S. (2011). Differential neural activity during search of specific and general autobiographical memories elicited by musical cues. *Neuropsychologia* 49, 2514–2526.
- Ford, J.H., Rubin, D.C., and Giovanello, K.S. (2016). The effects of song familiarity and age on phenomenological characteristics and neural recruitment during autobiographical memory retrieval. *Psychomusicology: Music, Mind, and Brain* 26, 199–210.
- Fóti, V.M. (2000). Descartes' intellectual and corporeal memories. In *Descartes' Natural Philosophy*, (Routledge), p.
- Foudil, S.-A., Pleche, C., and Macaluso, E. (2021). Memory for spatio-temporal contextual details during the retrieval of naturalistic episodes. *Sci Rep* 11, 14577.

- Freud, E., Stajduhar, A., Rosenbaum, R.S., Avidan, G., and Ganel, T. (2020). The COVID-19 pandemic masks the way people perceive faces. *Sci Rep* *10*, 22344.
- Frithsen, A., and Miller, M.B. (2014). The posterior parietal cortex: Comparing remember/know and source memory tests of recollection and familiarity. *Neuropsychologia* *61*, 31–44.
- Fritz, T., Jentschke, S., Gosselin, N., Sammler, D., Peretz, I., Turner, R., Friederici, A.D., and Koelsch, S. (2009). Universal Recognition of Three Basic Emotions in Music. *Current Biology* *19*, 573–576.
- Fuentes-Sánchez, N., Pastor, R., Escrig, M.A., Elipe-Miravet, M., and Pastor, M.C. (2021). Emotion elicitation during music listening: Subjective self-reports, facial expression, and autonomic reactivity. *Psychophysiology* *58*, e13884.
- Gazzaniga, M.S., Ivry, R.B., and Mangun, G.R. (2019). *Cognitive neuroscience: the biology of the mind* (New York: W.W. Norton & Company).
- Gerhold, K.A., and Bautista, D.M. (2009). Molecular and Cellular Mechanisms of Trigeminal Chemosensation. *Annals of the New York Academy of Sciences* *1170*, 184–189.
- Gilbert, K., and Schleuder, J. (1990). Effects of Color and Complexity in Still Photographs on Mental Effort and Memory. *Journalism Quarterly* *67*, 749–756.
- Glachet, O., and El Haj, M. (2021). Odor is more effective than a visual cue or a verbal cue for the recovery of autobiographical memories in AD. *Journal of Clinical and Experimental Neuropsychology* *43*, 129–143.
- Goddard, L., Pring, L., and Felmingham, N. (2005). The effects of cue modality on the quality of personal memories retrieved. *Memory* *13*, 79–86.
- Godden, D.R., and Baddeley, A.D. (1975). Context-Dependent Memory in Two Natural Environments: On Land and Underwater. *British Journal of Psychology* *66*, 325–331.
- Gold, B.P., Pearce, M.T., Mas-Herrero, E., Dagher, A., and Zatorre, R.J. (2019). Predictability and Uncertainty in the Pleasure of Music: A Reward for Learning? *J. Neurosci.* *39*, 9397–9409.
- Gold, J.J., Smith, C.N., Bayley, P.J., Shrager, Y., Brewer, J.B., Stark, C.E.L., Hopkins, R.O., and Squire, L.R. (2006). Item memory, source memory, and the medial temporal lobe: Concordant findings from fMRI and memory-impaired patients. *PNAS* *103*, 9351–9356.
- Goldberg, E.M., Wang, K., Goldberg, J., and Aliani, M. (2018). Factors affecting the ortho- and retronasal perception of flavors: A review. *Critical Reviews in Food Science and Nutrition* *58*, 913–923.
- Goodrich-Hunsaker, N.J., Gilbert, P.E., and Hopkins, R.O. (2009). The Role of the Human Hippocampus in Odor–Place Associative Memory. *Chemical Senses* *34*, 513–521.
- Gorgolewski, K., Burns, C., Madison, C., Clark, D., Halchenko, Y., Waskom, M., and Ghosh, S. (2011). Nipype: A Flexible, Lightweight and Extensible Neuroimaging Data Processing Framework in Python. *Frontiers in Neuroinformatics* *5*, 13.
- Gottfried, J.A. (2010). Central mechanisms of odour object perception. *Nat Rev Neurosci* *11*, 628–641.

- Graham, K.S., Simons, J.S., Pratt, K.H., Patterson, K., and Hodges, J.R. (2000). Insights from semantic dementia on the relationship between episodic and semantic memory. *Neuropsychologia* 38, 313–324.
- Greenberg, D., Eacott, M., Brechin, D., and Rubin, D. (2005a). Visual memory loss and autobiographical amnesia: a case study. *Neuropsychologia* 43, 1493–1502.
- Greenberg, D.L., Rice, H.J., Cooper, J.J., Cabeza, R., Rubin, D.C., and LaBar, K.S. (2005b). Co-activation of the amygdala, hippocampus and inferior frontal gyrus during autobiographical memory retrieval. *Neuropsychologia* 43, 659–674.
- Greve, D.N., and Fischl, B. (2009). Accurate and robust brain image alignment using boundary-based registration. *NeuroImage* 48, 63–72.
- Groussard, M., Chan, T.G., Coppalle, R., and Platel, H. (2019). Preservation of Musical Memory Throughout the Progression of Alzheimer’s Disease? Toward a Reconciliation of Theoretical, Clinical, and Neuroimaging Evidence. *JAD* 68, 857–883.
- Hackländer, R.P.M., Janssen, S.M.J., and Bermeitinger, C. (2018). An in-depth review of the methods, findings, and theories associated with odor-evoked autobiographical memory. *Psychonomic Bulletin & Review*.
- Haxby, J.V., Hoffman, E.A., and Gobbini, M.I. (2000). The distributed human neural system for face perception. *Trends in Cognitive Sciences* 4, 223–233.
- Hayes, J.P., VanElzakker, M.B., and Shin, L.M. (2012). Emotion and cognition interactions in PTSD: a review of neurocognitive and neuroimaging studies. *Front Integr Neurosci* 6, 89.
- Heilig, M.L. (1962). Sensorama Simulator.
- Henke, K. (2010). A model for memory systems based on processing modes rather than consciousness. *Nat Rev Neurosci* 11, 523–532.
- Herff, S.A., and Czernochowski, D. (2019). The role of divided attention and expertise in melody recognition. *Musicae Scientiae* 23, 69–86.
- Herz, R.S. (1998). Are odors the best cues to memory? A cross-modal comparison of associative memory stimuli. *Ann. N. Y. Acad. Sci.* 855, 670–674.
- Herz, R.S. (2004). A Naturalistic Analysis of Autobiographical Memories Triggered by Olfactory Visual and Auditory Stimuli. *Chem Senses* 29, 217–224.
- Herz, R.S. (2005). Odor-associative Learning and Emotion: Effects on Perception and Behavior. *Chemical Senses* 30, i250–i251.
- Herz, R.S. (2009). Aromatherapy Facts and Fictions: A Scientific Analysis of Olfactory Effects on Mood, Physiology and Behavior. *International Journal of Neuroscience* 119, 263–290.
- Herz, R., and Cupchik, G. (1995a). The Emotional Distinctiveness of Odor-evoked Memories. *Chemical Senses* 20, 517–528.
- Herz, R., and Cupchik, G. (1995b). The Emotional Distinctiveness of Odor-evoked Memories. *Chemical Senses* 20, 517–528.
- Herz, R., and Engen, T. (1996a). Odor memory: Review and analysis. *Psychonomic Bulletin & Review* 3, 300–313.

- Herz, R., and Engen, T. (1996b). Odor memory: Review and analysis. *Psychonomic Bulletin & Review* 3, 300–313.
- Herz, R.S., and Schooler, J.W. (2002). A naturalistic study of autobiographical memories evoked by olfactory and visual cues: testing the Proustian hypothesis. *American Journal of Psychology* 115, 21–32.
- Herz, R.S., Beland, S.L., and Hellerstein, M. (2004a). Changing Odor Hedonic Perception Through Emotional Associations in Humans. *International Journal of Comparative Psychology* 17, 315–338.
- Herz, R.S., Eliassen, J., Beland, S., and Souza, T. (2004b). Neuroimaging evidence for the emotional potency of odor-evoked memory. *Neuropsychologia* 42, 371–378.
- Hinton, P.B., and Henley, T.B. (1993). Cognitive and affective components of stimuli presented in three modes. *Bull. Psychon. Soc.* 31, 595–598.
- Holland, A.C., and Kensinger, E.A. (2010). Emotion and Autobiographical Memory. *Phys Life Rev* 7, 88–131.
- Holland, S.M., and Smulders, T.V. (2011). Do humans use episodic memory to solve a What-Where-When memory task? *Anim Cogn* 14, 95–102.
- Honing, H., ten Cate, C., Peretz, I., and Trehub, S.E. (2015). Without it no music: cognition, biology and evolution of musicality. *Philosophical Transactions of the Royal Society B: Biological Sciences* 370, 20140088.
- Hugenberg, K., Young, S.G., Sacco, D.F., and Bernstein, M.J. (2011). Social Categorization Influences Face Perception and Face Memory.
- Hummel, T., Iannilli, E., Frasnelli, J., Boyle, J., and Gerber, J. (2009). Central Processing of Trigeminal Activation in Humans. *Annals of the New York Academy of Sciences* 1170, 190–195.
- Hummel, T., Whitcroft, K.L., Andrews, P., Altundag, A., Cinghi, C., Costanzo, R.M., Damm, M., Frasnelli, J., Gudziol, H., Gupta, N., et al. (2017). Position paper on olfactory dysfunction. *Rhin* 54, 1–30.
- Huntenburg, J.M. (2014). Evaluating nonlinear coregistration of BOLD EPI and T1w images. Freie Universität Berlin.
- Huron, D. (2006). *Sweet anticipation: Music and the psychology of expectation* (Cambridge, MA, US: The MIT Press).
- Huron, D., and Margulis, E.H. (2010). Musical Expectancy and Thrills. In *Handbook of Music and Emotion: Theory, Research, Applications*, (Oxford: Oxford University Press), p.
- Husserl, E., and Gibson, W.R.B. (1931). *Ideas: general introduction to pure phenomenology* (London; New York: G. Allen & Unwin ; Macmillan Co.).
- Hutmacher, F. (2021). Do you remember? Similarities and differences between the earliest childhood memories for the five senses. *Memory* 29, 345–352.
- Jakubowski, K., and Ghosh, A. (2021). Music-evoked autobiographical memories in everyday life. *Psychology of Music* 49, 649–666.

- Jakubowski, K., Eerola, T., Tillmann, B., Perrin, F., and Heine, L. (2020). A Cross-Sectional Study of Reminiscence Bumps for Music-Related Memories in Adulthood. *Music & Science* 3, 205920432096505.
- Jakubowski, K., Belfi, A., and Eerola, T. (2021). Phenomenological Differences in Music- and Television-Evoked Autobiographical Memories. *Music Perception* 38, 435–455.
- James, W., Drummond, R., and Henry Holt and Company (1890). *The principles of psychology*.
- Janata, P. (2009a). The neural architecture of music-evoked autobiographical memories. *Cereb Cortex* 19, 2579–2594.
- Janata, P. (2009b). Music and the self. In *Music That Works*, R. Haas, and V. Brandes, eds. p.
- Janata, P. (2015). Chapter 11 - Neural basis of music perception. In *Handbook of Clinical Neurology*, M.J. Aminoff, F. Boller, and D.F. Swaab, eds. (Elsevier), pp. 187–205.
- Janata, P., Tomic, S.T., and Rakowski, S.K. (2007). Characterisation of music-evoked autobiographical memories. *Memory* 15, 845–860.
- JASP Team (2020). JASP (Version 0.14)[Computer software].
- Jellinek, J.S. (2004). Proust Remembered: Has Proust's Account of Odor-cued Autobiographical Memory Recall Really been Investigated? *Chemical Senses* 29, 455–458.
- Jenkinson, M., Bannister, P., Brady, M., and Smith, S. (2002). Improved Optimization for the Robust and Accurate Linear Registration and Motion Correction of Brain Images. *NeuroImage* 17, 825–841.
- Jiang, T., Soussignan, R., Rigaud, D., Martin, S., Royet, J.P., Brondel, L., and Schaal, B. (2008). Alliesthesia to food cues: heterogeneity across stimuli and sensory modalities. *Physiol Behav* 95, 464–470.
- Johansson, M., Mecklinger, A., and Treese, A.-C. (2004). Recognition memory for emotional and neutral faces: An event-related potential study. *Journal of Cognitive Neuroscience* 16, 1840–1853.
- Johnson, M.H. (2005). Subcortical face processing. *Nat Rev Neurosci* 6, 766–774.
- Johnston, R.A., and Edmonds, A.J. (2009). Familiar and unfamiliar face recognition: A review. *Memory* 17, 577–596.
- Jraissati, Y., and Deroy, O. (2021). Categorizing Smells: A Localist Approach. *Cognitive Science* 45, e12930.
- Juslin, P. (2013a). What does music express? Basic emotions and beyond. *Frontiers in Psychology* 4, 596.
- Juslin, P.N. (2013b). From everyday emotions to aesthetic emotions: Towards a unified theory of musical emotions. *Physics of Life Reviews* 10, 235–266.
- Kanwisher, N., McDermott, J., and Chun, M.M. (1997). The Fusiform Face Area: A Module in Human Extrastriate Cortex Specialized for Face Perception. *J. Neurosci.* 17, 4302–4311.

- Karlsson, K., Sikström, S., and Willander, J. (2013). The Semantic Representation of Event Information Depends on the Cue Modality: An Instance of Meaning-Based Retrieval. *PLOS ONE* *8*, e73378.
- Kätsyri, J. (2018). Those Virtual People all Look the Same to me: Computer-Rendered Faces Elicit a Higher False Alarm Rate Than Real Human Faces in a Recognition Memory Task. *Front Psychol* *9*.
- Kegel, L.C., Brugger, P., Frühholz, S., Grunwald, T., Hilfiker, P., Kohnen, O., Loertscher, M.L., Mersch, D., Rey, A., Sollfrank, T., et al. (2020). Dynamic human and avatar facial expressions elicit differential brain responses. *Social Cognitive and Affective Neuroscience* *15*, 347–358.
- Keightley, M.L., Chiew, K.S., Anderson, J.A.E., and Grady, C.L. (2011). Neural correlates of recognition memory for emotional faces and scenes. *Soc Cogn Affect Neurosci* *6*, 24–37.
- Keller, A., Zhuang, H., Chi, Q., Vosshall, L.B., and Matsunami, H. (2007). Genetic variation in a human odorant receptor alters odour perception. *Nature* *449*, 468–472.
- Khan, R.M., Luk, C.-H., Flinker, A., Aggarwal, A., Lapid, H., Haddad, R., and Sobel, N. (2007). Predicting Odor Pleasantness from Odorant Structure: Pleasantness as a Reflection of the Physical World. *Journal of Neuroscience* *27*, 10015–10023.
- Klein, A., Ghosh, S.S., Bao, F.S., Giard, J., Häme, Y., Stavsky, E., Lee, N., Rossa, B., Reuter, M., Neto, E.C., et al. (2017). Mindboggling morphometry of human brains. *PLOS Computational Biology* *13*, e1005350.
- Knez, I., Ljunglöf, L., Arshamian, A., and Willander, J. (2017). Self-grounding visual, auditory and olfactory autobiographical memories. *Consciousness and Cognition* *52*, 1–8.
- Koelsch, S. (2018). Investigating the Neural Encoding of Emotion with Music. *Neuron* *98*, 1075–1079.
- Koelsch, S., Jacobs, A.M., Menninghaus, W., Liebal, K., Klann-Delius, G., von Scheve, C., and Gebauer, G. (2015). The quartet theory of human emotions: An integrative and neurofunctional model. *Physics of Life Reviews* *13*, 1–27.
- Kontaris, I., East, B.S., and Wilson, D.A. (2020). Behavioral and Neurobiological Convergence of Odor, Mood and Emotion: A Review. *Front. Behav. Neurosci.* *14*.
- Korsakoff, S. (1889). Etude médico-psychologique sur une forme des maladies de la mémoire. *Revue Philosophique de La France et de l'Étranger* *28*, 501–530.
- Krumhansl, C.L. (2017). Listening Niches across a Century of Popular Music. *Front Psychol* *8*.
- Krumhansl, C.L., and Zupnick, J.A. (2013). Cascading Reminiscence Bumps in Popular Music. *Psychol Sci* *24*, 2057–2068.
- La Corte, V., Sperduti, M., Abichou, K., and Piolino, P. (2019). Episodic Memory Assessment and Remediation in Normal and Pathological Aging Using Virtual Reality: A Mini Review. *Front. Psychol.* *0*.
- Lanczos, C. (1964). Evaluation of Noisy Data. *Journal of the Society for Industrial and Applied Mathematics Series B Numerical Analysis* *1*, 76–85.
- Lane, G., Zhou, G., Noto, T., and Zelano, C. (2020). Assessment of direct knowledge of the human olfactory system. *Exp Neurol* *329*, 113304.

- Laney, C., Campbell, H.V., Heuer, F., and Reisberg, D. (2004). Memory for thematically arousing events. *Memory & Cognition* 32, 1149–1159.
- Lange, K., and Czernochowski, D. (2013). Does this sound familiar? Effects of timbre change on episodic retrieval of novel melodies. *Acta Psychologica* 143, 136–145.
- Larsson, M., Blåvarg, E.C., and Jönsson, F. (2009). Bad Odors Stick Better Than Good Ones: Olfactory Qualities and Odor Recognition. *Experimental Psychology* 56, 375–380.
- Larsson, M., Willander, J., Karlsson, K., and Arshamian, A. (2014). Olfactory LOVER: behavioral and neural correlates of autobiographical odor memory. *Front. Psychol.* 5.
- Lawless, H.T. (1978). Recognition of common odors, pictures, and simple shapes. *Perception & Psychophysics* 24, 493–495.
- Lehn, H., Kjønigsen, L.J., Kjelvik, G., and Håberg, A.K. (2013). Hippocampal involvement in retrieval of odor vs. object memories. *Hippocampus* 23, 122–128.
- Levine, B., Svoboda, E., Hay, J.F., Winocur, G., and Moscovitch, M. (2002). Aging and autobiographical memory: Dissociating episodic from semantic retrieval. *Psychology and Aging* 17, 677–689.
- Levy, D.A. (2004). Impaired odor recognition memory in patients with hippocampal lesions. *Learning & Memory* 11, 794–796.
- Lima, C.F., Brancatisano, O., Fancourt, A., Müllensiefen, D., Scott, S.K., Warren, J.D., and Stewart, L. (2016). Impaired socio-emotional processing in a developmental music disorder. *Sci Rep* 6, 34911.
- Lindquist, K.A., Wager, T.D., Kober, H., Bliss-Moreau, E., and Barrett, L.F. (2012). The brain basis of emotion: A meta-analytic review. *Behavioral and Brain Sciences* 35, 121–143.
- Liu, C.H., Chen, W., and Ward, J. (2014). Remembering faces with emotional expressions. *Frontiers in Psychology* 5, 1439.
- Livermore, A., and Laing, D.G. (1996). Influence of training and experience on the perception of multicomponent odor mixtures. *J Exp Psychol Hum Percept Perform* 22, 267–277.
- Longmore, C.A., Liu, C.H., and Young, A.W. (2008). Learning faces from photographs. *Journal of Experimental Psychology: Human Perception and Performance* 34, 77–100.
- Lundqvist, L.-O., Carlsson, F., Hilmersson, P., and Juslin, P.N. (2009). Emotional responses to music: experience, expression, and physiology. *Psychology of Music* 37, 61–90.
- Madan, C.R. (2020). Rethinking the definition of episodic memory. *Can J Exp Psychol* 74, 183–192.
- Madigan, S., and O'Hara, R. (1992). Short-term memory at the turn of the century: Mary Whiton Calkins's memory research. *American Psychologist* 47, 170–174.
- Majid, A. (2021a). Human Olfaction at the Intersection of Language, Culture, and Biology. *Trends in Cognitive Sciences* 25, 111–123.
- Majid, A. (2021b). Olfactory Language Requires an Integrative and Interdisciplinary Approach. *Trends in Cognitive Sciences* 25, 421–422.

- Majid, A., and Burenhult, N. (2014). Odors are expressible in language, as long as you speak the right language. *Cognition* 130, 266–270.
- Mandaïron, N., Poncelet, J., Bensafi, M., and Didier, A. (2009). Humans and Mice Express Similar Olfactory Preferences. *PLoS One* 4, e4209.
- Manesse, C., Ferdenzi, C., Mantel, M., Sabri, M., Bessy, M., Fournel, A., Faure, F., Bellil, D., Landis, B.M., 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 Quality and Preference* 93, 104273.
- Mantel, M., Ferdenzi, C., Roy, J.-M., and 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 Topogr* 32, 977–986.
- Mao, X., You, Y., Li, W., and Guo, C. (2015). Emotion impairs extrinsic source memory—An ERP study. *Biological Psychology* 110, 182–189.
- Maquet, P. (2001). The role of sleep in learning and memory. *Science* 294, 1048–1052.
- Masaoka, Y., Sugiyama, H., Katayama, A., Kashiwagi, M., and Homma, I. (2012). Slow Breathing and Emotions Associated with Odor-Induced Autobiographical Memories. *Chemical Senses* 37, 379–388.
- Masaoka, Y., Izumizaki, M., and Homma, I. (2014). Where is the rhythm generator for emotional breathing? *Prog Brain Res* 209, 367–377.
- Masaoka, Y., Sugiyama, H., Yoshida, M., Yoshikawa, A., Honma, M., Koiwa, N., Kamijo, S., Watanabe, K., Kubota, S., Iizuka, N., et al. (2021). Odors Associated With Autobiographical Memory Induce Visual Imagination of Emotional Scenes as Well as Orbitofrontal-Fusiform Activation. *Frontiers in Neuroscience* 15, 974.
- Mas-Herrero, E., Zatorre, R.J., Rodriguez-Fornells, A., and Marco-Pallarés, J. (2014). Dissociation between Musical and Monetary Reward Responses in Specific Musical Anhedonia. *Current Biology* 24, 699–704.
- Matsumoto, D., Keltner, D., Shiota, M.N., O’Sullivan, M., and Frank, M. (2008). Facial expressions of emotion. In *Handbook of Emotions*, 3rd Ed, (New York, NY, US: The Guilford Press), pp. 211–234.
- Maurer, D., Grand, R.L., and Mondloch, C.J. (2002). The many faces of configural processing. *Trends in Cognitive Sciences* 6, 255–260.
- McDermott, K.B., Szpunar, K.K., and Christ, S.E. (2009). Laboratory-based and autobiographical retrieval tasks differ substantially in their neural substrates. *Neuropsychologia* 47, 2290–2298.
- McGann, J.P. (2017). Poor human olfaction is a 19th-century myth. *Science* 356, eaam7263.
- Meister, M. (2015). On the dimensionality of odor space. *Elife* 4, e07865.
- Melton, A.W. (1963). Implications of short-term memory for a general theory of memory. *Journal of Verbal Learning and Verbal Behavior* 2, 1–21.

- Meunier, D., Fonlupt, P., Saive, A.-L., Plailly, J., Ravel, N., and Royet, J.-P. (2014). Modular structure of functional networks in olfactory memory. *NeuroImage* *95*, 264–275.
- Meyer, L.B. (1956). *Emotion and meaning in music* (University of Chicago Press: Chicago).
- Miles, A.N., and Berntsen, D. (2011). Odour-induced mental time travel into the past and future: Do odour cues retain a unique link to our distant past? *Memory* *19*, 930–940.
- Mitchell, K.J., and Johnson, M.K. (2009). Source monitoring 15 years later: What have we learned from fMRI about the neural mechanisms of source memory? *Psychological Bulletin* *135*, 638–677.
- Molnar-Szakacs, I., and Overy, K. (2006). Music and mirror neurons: from motion to 'e'motion. *Soc Cogn Affect Neurosci* *1*, 235–241.
- Monge, Z.A., Wing, E.A., Stokes, J., and Cabeza, R. (2018). Search and recovery of autobiographical and laboratory memories: Shared and distinct neural components. *Neuropsychologia* *110*, 44–54.
- Moss, A., Miles, C., Elsley, J., and Johnson, A.J. (2019). Olfactory working memory: exploring the differences in n-back memory for high and low verbalisable odorants. *Memory* *27*, 1319–1344.
- Mukudi, P.B.L., and Hills, P.J. (2019). The combined influence of the own-age, -gender, and -ethnicity biases on face recognition. *Acta Psychologica* *194*, 1–6.
- Murphy, C., Cain, W.S., Gilmore, M.M., and Skinner, R.B. (1991). Sensory and Semantic Factors in Recognition Memory for Odors and Graphic Stimuli: Elderly versus Young Persons. *The American Journal of Psychology* *104*, 161.
- Nader, K., Schafe, G.E., and Le Doux, J.E. (2000). Fear memories require protein synthesis in the amygdala for reconsolidation after retrieval. *Nature* *406*, 722–726.
- Narme, P., Peretz, I., Strub, M.-L., and Ergis, A.-M. (2016). Emotion effects on implicit and explicit musical memory in normal aging. *Psychology and Aging* *31*, 902–913.
- Nietzsche, F. (1908). *Maximes et pointes*. In *Le Crépuscule Des Idoles Ou Comment on Philosophe Au Marteau*, (Paris: Mercure de France), p.
- Nineuil, C., Dellacherie, D., and Samson, S. (2020). The Impact of Emotion on Musical Long-Term Memory. *Front. Psychol.* *11*, 2110.
- Olofsson, J.K., and Gottfried, J.A. (2015). The muted sense: neurocognitive limitations of olfactory language. *Trends in Cognitive Sciences* *19*, 314–321.
- Olofsson, J.K., and Pierzchajlo, S. (2021). Olfactory Language: Context Is Everything. *Trends in Cognitive Sciences* *25*, 419–420.
- Olsson, M., Lundgren, E., Soares, S., and Johansson, M. (2009). Odor Memory Performance and Memory Awareness: A Comparison to Word Memory Across Orienting Tasks and Retention Intervals. *Chemosensory Perception* *2*, 161–171.
- Palombo, D.J., Sheldon, S., and Levine, B. (2018). Individual Differences in Autobiographical Memory. *TRENDS COGN. SCI.* *22*, 583–597.

- Panksepp, J. (1995). The Emotional Sources of Chills Induced by Music. *Music Perception* 13, 171–207.
- Pascalis, O., and Kelly, D.J. (2009). The Origins of Face Processing in Humans: Phylogeny and Ontogeny. *Perspect Psychol Sci* 4, 200–209.
- Pause, B., Zlomuzica, A., Kinugawa, K., Mariani, J., Pietrowsky, R., and Dere, E. (2013). Perspectives on Episodic-Like and Episodic Memory. *Frontiers in Behavioral Neuroscience* 7, 33.
- Pearce, M.T. (2018). Statistical learning and probabilistic prediction in music cognition: mechanisms of stylistic enculturation. *Annals of the New York Academy of Sciences* 1423, 378–395.
- Peretz, I., and Coltheart, M. (2003). Modularity of music processing. *Nat Neurosci* 6, 688–691.
- Peretz, I., and Zatorre, R.J. (2005). Brain Organization for Music Processing. *Annu. Rev. Psychol.* 56, 89–114.
- Peretz, I., Gosselin, N., Belin, P., Zatorre, R.J., Plailly, J., and Tillmann, B. (2009). Music lexical networks: the cortical organization of music recognition. *Ann N Y Acad Sci* 1169, 256–265.
- Peynircioğlu, Z.F., Mungan, E., and Schwartz, B.L. (2018). Chapter 19 - Challenges in Music Memory Research. In *The Handbook of Research Methods in Human Memory*, (Routledge Handbooks Online), p.
- Phelps, E.A., and Anderson, A.K. (1997). Emotional memory: what does the amygdala do? *Curr Biol* 7, R311-4.
- Phelps, E.A., and Sharot, T. (2008). How (and Why) Emotion Enhances the Subjective Sense of Recollection. *Curr Dir Psychol Sci* 17, 147–152.
- Picard, L., Eustache, F., and Piolino, P. (2009). De la mémoire épisodique à la mémoire autobiographique : approche développementale. *L'Année psychologique* 109, 197–236.
- Pinker, S. (1998). *How the Mind Works* (London: Allen Lane).
- Piolino, P., Desgranges, B., Benali, K., and Eustache, F. (2002). Episodic and semantic remote autobiographical memory in ageing. *Memory* 10, 239–257.
- Piolino, P., Desgranges, B., Clarys, D., Guillery-Girard, B., Taconnat, L., Isingrini, M., and Eustache, F. (2006). Autobiographical memory, auto-noetic consciousness, and self-perspective in aging. *Psychology and Aging* 21, 510–525.
- Piolino, P., Desgranges, B., and Eustache, F. (2009). Episodic autobiographical memories over the course of time: Cognitive, neuropsychological and neuroimaging findings. *Neuropsychologia* 47, 2314–2329.
- Plailly, J., Tillmann, B., and Royet, J.-P. (2007). The Feeling of Familiarity of Music and Odors: The Same Neural Signature? *Cerebral Cortex* 17, 2650–2658.
- Plailly, J., Luangraj, N., Nicklaus, S., Issanchou, S., Royet, J.P., and Sulmont-Rosse, C. (2011). Alliesthesia is greater for odors of fatty foods than of non-fat foods. *Appetite* 57, 615–622.

- Plancher, G., and Barrouillet, P. (2020). On some of the main criticisms of the modal model: Reappraisal from a TBRS perspective. *Mem Cogn* 48, 455–468.
- Plancher, G., and Piolino, P. (2017). Virtual Reality for Assessment of Episodic Memory in Normal and Pathological Aging. In *Virtual Reality for Assessment of Episodic Memory in Normal and Pathological Aging*, (Oxford University Press), p.
- Plancher, G., Barra, J., Orriols, E., and Piolino, P. (2013). The influence of action on episodic memory: A virtual reality study. *Quarterly Journal of Experimental Psychology* 66, 895–909.
- Platel, H., and Groussard, M. (2010). La mémoire sémantique musicale : apport des données de la neuropsychologie clinique et de la neuro-imagerie fonctionnelle. *Revue de neuropsychologie Volume 2*, 61–69.
- Platel, H., Baron, J.-C., Desgranges, B., Bernard, F., and Eustache, F. (2003). Semantic and episodic memory of music are subserved by distinct neural networks. *NeuroImage* 20, 244–256.
- Poppenk, J., Evensmoen, H.R., Moscovitch, M., and Nadel, L. (2013). Long-axis specialization of the human hippocampus. *Trends in Cognitive Sciences* 17, 230–240.
- Power, J.D., Mitra, A., Laumann, T.O., Snyder, A.Z., Schlaggar, B.L., and Petersen, S.E. (2014). Methods to detect, characterize, and remove motion artifact in resting state fMRI. *NeuroImage* 84, 320–341.
- Pralus, A., Fornoni, L., Bouet, R., Gomot, M., Bhatara, A., Tillmann, B., and Caclin, A. (2019). Emotional prosody in congenital amusia: Impaired and spared processes. *Neuropsychologia* 134, 107234.
- Proust, M. (1913). *Du côté de chez Swann*.
- Pützer, A., Otto, T., and Wolf, O. (2019). Odors Are More Sensitive to Evaluative Conditioning than Sounds. *Chemosensory Perception*.
- Rapcsak, S.Z. (2013). Face Recognition. *The Roots of Cognitive Neuroscience: Behavioral Neurology and Neuropsychology*.
- Rapcsak, S.Z. (2019). Face Recognition. *Curr Neurol Neurosci Rep* 19, 41.
- Rapcsak, S.Z., and Edmonds, E.C. (2011). The Executive Control of Face Memory. *Behavioural Neurology* 24, 285–298.
- Reagh, Z.M., and Ranganath, C. (2018). What does the functional organization of cortico-hippocampal networks tell us about the functional organization of memory? *Neuroscience Letters* 680, 69–76.
- Reisert, J., and Restrepo, D. (2009). Molecular Tuning of Odorant Receptors and Its Implication for Odor Signal Processing. *Chem Senses* 34, 535–545.
- Renoult, L., and Rugg, M.D. (2020). An historical perspective on Endel Tulving’s episodic-semantic distinction. *Neuropsychologia* 139.
- Renoult, L., Irish, M., Moscovitch, M., and Rugg, M.D. (2019). From Knowing to Remembering: The Semantic–Episodic Distinction. *Trends in Cognitive Sciences* 23, 1041–1057.

- Rey, L., Désoche, C., Thévenet, M., Garcia, S., Tillmann, B., and Plailly, J. (b). Episodic memory evoked by odor compared to musical excerpts and faces: Modality-specific effect of wanting, but modality unspecific effect of valence.
- Rey, L., Désoche, C., Thévenet, M., Garcia, S., Tillmann, B., and Plailly, J. (a). Episodic memory for odors, music and faces – an ecological laboratory approach using virtual reality.
- Ribot, T. (1870). *La psychologie anglaise contemporaine (école expérimentale)* (Ladrange).
- Richardson, J.T., and Zucco, G.M. (1989). Cognition and olfaction: a review. *Psychol Bull* 105, 352–360.
- Righi, S., Marzi, T., Toscani, M., Baldassi, S., Ottonello, S., and Viggiano, M.P. (2012). Fearful expressions enhance recognition memory: Electrophysiological evidence. *Acta Psychologica* 139, 7–18.
- Ritchey, M., and Cooper, R.A. (2020). Deconstructing the Posterior Medial Episodic Network. *Trends in Cognitive Sciences* 24, 451–465.
- Ritchey, M., Libby, L.A., and Ranganath, C. (2015). Chapter 3 - Cortico-hippocampal systems involved in memory and cognition: the PMAT framework. In *Progress in Brain Research*, S. O'Mara, and M. Tsanov, eds. (Elsevier), pp. 45–64.
- Roediger, H.L., and McDermott, K.B. (2013). Two types of event memory (commentary). *Proceedings of the National Academy of Sciences* 110, 20856–20857.
- Rossion, B. (2018). Humans Are Visual Experts at Unfamiliar Face Recognition. *Trends in Cognitive Sciences* 22, 471–472.
- Royet, J.-P., Plailly, J., Delon-Martin, C., Kareken, D.A., and Segebarth, C. (2003). fMRI of emotional responses to odors: influence of hedonic valence and judgment, handedness, and gender. *NeuroImage* 20, 713–728.
- Royet, J.-P., Morin-Audebrand, L., Cerf-Ducastel, B., Haase, L., Issanchou, S., Murphy, C., Fonlupt, P., Sulmont-Rossé, C., and Plailly, J. (2011). True and False Recognition Memories of Odors Induce Distinct Neural Signatures. *Frontiers in Human Neuroscience* 5, 65.
- Royet, J.-P., Delon-Martin, C., and Plailly, J. (2013). Odor mental imagery in non-experts in odors: a paradox? *Frontiers in Human Neuroscience* 7, 87.
- Royet, J.-P., Saive, A.-L., Plailly, J., and Veyrac, A. (2015). Être parfumeur, une question de prédisposition ou d'entraînement ? In *L'Art Olfactif Contemporain*, (Classiques Garnier), p.
- Rubin, D.C., and Berntsen, D. (2009). The Frequency of Voluntary and Involuntary Autobiographical Memories across the Lifespan. *Mem Cognit* 37, 679–688.
- Rubin, D.C., Groth, E., and Goldsmith, D.J. (1984). Olfactory Cuing of Autobiographical Memory. *The American Journal of Psychology* 97, 493.
- Rugg, M.D., and Vilberg, K.L. (2013). Brain networks underlying episodic memory retrieval. *Current Opinion in Neurobiology* 23, 255–260.
- Russell, R., Duchaine, B., and Nakayama, K. (2009). Super-recognizers: People with extraordinary face recognition ability. *Psychonomic Bulletin & Review* 16, 252–257.

Sabatinelli, D., Fortune, E.E., Li, Q., Siddiqui, A., Krafft, C., Oliver, W.T., Beck, S., and Jeffries, J. (2011). Emotional perception: Meta-analyses of face and natural scene processing. *NeuroImage* *54*, 2524–2533.

Saive, A.-L., Ravel, N., Thévenet, M., Royet, J.-P., and Plailly, J. (2013). A novel experimental approach to episodic memory in humans based on the privileged access of odors to memories. *Journal of Neuroscience Methods* *213*, 22–31.

Saive, A.-L., Royet, J.-P., Ravel, N., Thévenet, M., Garcia, S., and Plailly, J. (2014a). A unique memory process modulated by emotion underpins successful odor recognition and episodic retrieval in humans. *Frontiers in Behavioral Neuroscience* *8*.

Saive, A.-L., Royet, J.-P., and Plailly, J. (2014b). A review on the neural bases of episodic odor memory: from laboratory-based to autobiographical approaches. *Frontiers in Behavioral Neuroscience* *8*.

Saive, A.-L., Royet, J.-P., Garcia, S., Thévenet, M., and Plailly, J. (2015). “What-Where-Which” Episodic Retrieval Requires Conscious Recollection and Is Promoted by Semantic Knowledge. *PLOS ONE* *10*, e0143767.

Salakka, I., Pitkänieni, A., Pentikäinen, E., Mikkonen, K., Saari, P., Toiviainen, P., and Särkämö, T. (2021). What makes music memorable? Relationships between acoustic musical features and music-evoked emotions and memories in older adults. *PLOS ONE* *16*, e0251692.

Salimpoor, V.N., Benovoy, M., Longo, G., Cooperstock, J.R., and Zatorre, R.J. (2009). The Rewarding Aspects of Music Listening Are Related to Degree of Emotional Arousal. *PLoS ONE* *4*, e7487.

Samson, S., Dellacherie, D., and Platel, H. (2009). Emotional Power of Music in Patients with Memory Disorders: Clinical Implications of Cognitive Neuroscience. *Annals of the New York Academy of Sciences* *1169*, 245–255.

Särkämö, T., Tervaniemi, M., and Huutilainen, M. (2013). Music perception and cognition: development, neural basis, and rehabilitative use of music. *WIREs Cogn Sci* *4*, 441–451.

Sato, W., and Yoshikawa, S. (2013). Recognition Memory for Faces and Scenes. *The Journal of General Psychology* *140*, 1–15.

Satterthwaite, T.D., Elliott, M.A., Gerraty, R.T., Ruparel, K., Loughhead, J., Calkins, M.E., Eickhoff, S.B., Hakonarson, H., Gur, R.C., Gur, R.E., et al. (2013). An improved framework for confound regression and filtering for control of motion artifact in the preprocessing of resting-state functional connectivity data. *NeuroImage* *64*, 240–256.

Schaal, N.K., Javadi, A.-H., Halpern, A.R., Pollok, B., and Banissy, M.J. (2015). Right parietal cortex mediates recognition memory for melodies. *Eur. J. Neurosci.* *42*, 1660–1666.

Schacter, D.L., and Addis, D.R. (2007). The cognitive neuroscience of constructive memory: remembering the past and imagining the future. *Phil. Trans. R. Soc. B* *362*, 773–786.

Schacter, D.L., Addis, D.R., Hassabis, D., Martin, V.C., Spreng, R.N., and Szpunar, K.K. (2012). The Future of Memory: Remembering, Imagining, and the Brain. *Neuron* *76*, 677–694.

Schellenberg, E.G., and Habashi, P. (2015). Remembering the melody and timbre, forgetting the key and tempo. *Mem Cogn* *43*, 1021–1031.

- Schellenberg, E.G., Stalinski, S.M., and Marks, B.M. (2014). Memory for surface features of unfamiliar melodies: independent effects of changes in pitch and tempo. *Psychological Research* 78, 84–95.
- Schnitzspahn, K.M., Ihle, A., Henry, J.D., Rendell, P.G., and Kliegel, M. (2011). The age-prospective memory-paradox: an exploration of possible mechanisms. *International Psychogeriatrics* 23, 583–592.
- Schubert, E. (2013). Emotion felt by the listener and expressed by the music: literature review and theoretical perspectives. *Frontiers in Psychology* 4, 837.
- Schulkind, M.D. (2009). Is Memory for Music Special? *Annals of the New York Academy of Sciences* 1169, 216–224.
- Schulkind, M.D., Hennis, L.K., and Rubin, D.C. (1999). Music, emotion, and autobiographical memory: They're playing your song. *Mem Cogn* 27, 948–955.
- Scoville, W.B., and Milner, B. (1957). LOSS OF RECENT MEMORY AFTER BILATERAL HIPPOCAMPAL LESIONS. *Journal of Neurology, Neurosurgery & Psychiatry* 20, 11–21.
- Scully, I.D., Napper, L.E., and Hupbach, A. (2017). Does reactivation trigger episodic memory change? A meta-analysis. *Neurobiology of Learning and Memory* 142, 99–107.
- Sezille, C., Messaoudi, B., Bertrand, A., Jousain, P., Thévenet, M., and Bensafi, M. (2013). A portable experimental apparatus for human olfactory fMRI experiments. *J. Neurosci. Methods* 218, 29–38.
- Shallice, T., and Warrington, E.K. (1970). Independent Functioning of Verbal Memory Stores: A Neuropsychological Study. *Quarterly Journal of Experimental Psychology* 22, 261–273.
- Sharot, T., Delgado, M.R., and Phelps, E.A. (2004). How emotion enhances the feeling of remembering. *Nat Neurosci* 7, 1376–1380.
- Sheldon, S., and Donahue, J. (2017). More than a feeling: Emotional cues impact the access and experience of autobiographical memories. *Memory & Cognition* 45, 731–744.
- Sheldon, S., Fenerci, C., and Gurguryan, L. (2019). A Neurocognitive Perspective on the Forms and Functions of Autobiographical Memory Retrieval. *Front. Syst. Neurosci.* 0.
- Sheldon, S., Williams, K., Harrington, S., and Otto, A.R. (2020). Emotional cue effects on accessing and elaborating upon autobiographical memories. *Cognition* 198, 104217.
- Shimamura, A.P., Ross, J.G., and Bennett, H.D. (2006). Memory for facial expressions: The power of a smile. *Psychonomic Bulletin & Review* 13, 217–222.
- Silva, J.S.C.D., Barbosa, F.F., Fonsêca, É.K.G.D., Albuquerque, F.D.S., Cheke, L.G., and Fernández-Calvo, B. (2020). Load effect on what-where-when memory in younger and older adults. *Neuropsychol Dev Cogn B Aging Neuropsychol Cogn* 27, 841–853.
- Smith, S.A. (2019). Virtual reality in episodic memory research: A review. *Psychon Bull Rev* 26, 1213–1237.
- Smith, S.M., Glenberg, A., and Bjork, R.A. (1978). Environmental context and human memory. *Memory & Cognition* 6, 342–353.

- Squire, L.R. (1992). Memory and the hippocampus: a synthesis from findings with rats, monkeys, and humans. *Psychol Rev* 99, 195–231.
- Squire, L.R. (2004). Memory systems of the brain: A brief history and current perspective. *Neurobiology of Learning and Memory* 82, 171–177.
- Squire, L., and Zola-Morgan, S. (1991). The medial temporal lobe memory system. *Science* 253, 1380–1386.
- Squire, L.R., and Wixted, J.T. (2011). The Cognitive Neuroscience of Human Memory Since H.M. *Annu Rev Neurosci* 34, 259–288.
- Squire, L.R., and Zola-Morgan, S. (1988). Memory: brain systems and behavior. *Trends in Neurosciences* 11, 170–175.
- St Augustin d'Hippone (397AD). LIVRE X. In *Les Confessions*, p.
- St. Jacques, P.L., and Schacter, D.L. (2013). Modifying Memory: Selectively Enhancing and Updating Personal Memories for a Museum Tour by Reactivating Them. *Psychol Sci* 24, 537–543.
- St. Jacques, P.L., Kragel, P.A., and Rubin, D.C. (2011). Dynamic neural networks supporting memory retrieval. *NeuroImage* 57, 608–616.
- Stalinski, S.M., and Schellenberg, E.G. (2013). Listeners remember music they like. *Journal of Experimental Psychology: Learning, Memory, and Cognition* 39, 700–716.
- Staniloiu, A., Kordon, A., and Markowitsch, H.J. (2020). Quo vadis “episodic memory”? - Past, present, and perspective. *Neuropsychologia* 141, 107362.
- Stantic, M., Brewer, R., Duchaine, B., Banissy, M.J., Bate, S., Susilo, T., Catmur, C., and Bird, G. (2021). The Oxford Face Matching Test: A non-biased test of the full range of individual differences in face perception. *Behav Res.*
- Stevens, C.J. (2012). Music Perception and Cognition: A Review of Recent Cross-Cultural Research. *Topics in Cognitive Science* 4, 653–667.
- Stevenson, R.J. (2010). An Initial Evaluation of the Functions of Human Olfaction. *Chemical Senses* 35, 3–20.
- Stevenson, R.J., and Mahmut, M.K. (2013). The accessibility of semantic knowledge for odours that can and cannot be named. *The Quarterly Journal of Experimental Psychology* 66, 1414–1431.
- Stevenson, R.F., Reagh, Z.M., Chun, A.P., Murray, E.A., and Yassa, M.A. (2020). Pattern Separation and Source Memory Engage Distinct Hippocampal and Neocortical Regions during Retrieval. *J. Neurosci.* 40, 843–851.
- Stickgold, R. (2005). Sleep-dependent memory consolidation. *Nature* 437, 1272–1278.
- Sunday, M.A., and Gauthier, I. (2018). Face Expertise for Unfamiliar Faces: A Commentary on Young and Burton’s “Are We Face Experts?” 1, 7.
- Svoboda, E., McKinnon, M.C., and Levine, B. (2006). The functional neuroanatomy of autobiographical memory: A meta-analysis. *Neuropsychologia* 44, 2189–2208.

- Swarbrick, D., Seibt, B., Grinspun, N., and Vuoskoski, J.K. (2021). Corona Concerts: The Effect of Virtual Concert Characteristics on Social Connection and Kama Muta. *Front Psychol* *12*, 648448.
- Talamini, F., Altoè, G., Carretti, B., and Grassi, M. (2017). Musicians have better memory than nonmusicians: A meta-analysis. *PLoS ONE* *12*, e0186773.
- Tanner, W.P.Jr., and Swets, J.A. (1954). A decision-making theory of visual detection. *Psychological Review* *61*, 401–409.
- Taruffi, L., and Koelsch, S. (2014). The Paradox of Music-Evoked Sadness: An Online Survey. *PLOS ONE* *9*, e110490.
- Taubert, J., Apthorp, D., Aagten-Murphy, D., and Alais, D. (2011). The role of holistic processing in face perception: Evidence from the face inversion effect. *Vision Research* *51*, 1273–1278.
- Thompson, P. (1980). Margaret Thatcher: a new illusion. *Perception* *9*, 483–484.
- Tillmann, B., Albouy, P., and Caclin, A. (2015). Chapter 33 - Congenital amusias. In *Handbook of Clinical Neurology*, M.J. Aminoff, F. Boller, and D.F. Swaab, eds. (Elsevier), pp. 589–605.
- Toffolo, M.B.J., Smeets, M.A.M., and van den Hout, M.A. (2012). Proust revisited: odours as triggers of aversive memories. *Cogn Emot* *26*, 83–92.
- Toller, S.V. (1999). Assessing the Impact of Anosmia: Review of a Questionnaire's Findings. *Chemical Senses* *24*, 705–712.
- Treiber, J.M., White, N.S., Steed, T.C., Bartsch, H., Holland, D., Farid, N., McDonald, C.R., Carter, B.S., Dale, A.M., and Chen, C.C. (2016). Characterization and Correction of Geometric Distortions in 814 Diffusion Weighted Images. *PLOS ONE* *11*, e0152472.
- Trinkler, I., King, J.A., Doeller, C.F., Rugg, M.D., and Burgess, N. (2009). Neural bases of autobiographical support for episodic recollection of faces. *Hippocampus* *19*, 718–730.
- Tulving, E. (1972). Episodic and semantic memory. In *Organization of Memory*, (Oxford, England: Academic Press), pp. xiii, 423–xiii, 423.
- Tulving, E. (1982). Synergistic ephory in recall and recognition. *Canadian Journal of Psychology/Revue Canadienne de Psychologie* *36*, 130–147.
- Tulving, E. (1985). Memory and consciousness. *Canadian Psychology/Psychologie Canadienne* *26*, 1–12.
- Tulving, E. (1995). Organization of memory: Quo vadis? In *The Cognitive Neurosciences*, (Cambridge, MA, US: The MIT Press), pp. 839–853.
- Tulving, E. (2002). Episodic Memory: From Mind to Brain. *Annu. Rev. Psychol.* *53*, 1–25.
- Tustison, N.J., Avants, B.B., Cook, P.A., Zheng, Y., Egan, A., Yushkevich, P.A., and Gee, J.C. (2010). N4ITK: Improved N3 Bias Correction. *IEEE Transactions on Medical Imaging* *29*, 1310–1320.
- Vermetten, E., and Bremner, J.D. (2003). Olfaction as a traumatic reminder in posttraumatic stress disorder: case reports and review. *J Clin Psychiatry* *64*, 202–207.

- Veyrac, A., Allerborn, M., Gros, A., Michon, F., Raguet, L., Kenney, J., Godinot, F., Thevenet, M., García, S., Messaoudi, B., et al. (2015). Memory of Occasional Events in Rats: Individual Episodic Memory Profiles, Flexibility, and Neural Substrate. *The Journal of Neuroscience : The Official Journal of the Society for Neuroscience* *35*, 7575–7586.
- Vieillard, S., and Gilet, A.-L. (2013). Age-related differences in affective responses to and memory for emotions conveyed by music: a cross-sectional study. *Frontiers in Psychology* *4*, 711.
- Vieillard, S., Peretz, I., Gosselin, N., Khalfa, S., Gagnon, L., and Bouchard, B. (2008). Happy, sad, scary and peaceful musical excerpts for research on emotions. *Cognition & Emotion* *22*, 720–752.
- Vieillard, S., Ronat, L., Baccarani, A., Schaal, B., Baudouin, J.-Y., and Brochard, R. (2021). Age differences in olfactory affective responses: evidence for a positivity effect and an emotional dedifferentiation. *Aging, Neuropsychology, and Cognition* *28*, 570–583.
- Vuilleumier, P., and Pourtois, G. (2007). Distributed and interactive brain mechanisms during emotion face perception: Evidence from functional neuroimaging. *Neuropsychologia* *45*, 174–194.
- Vuilleumier, P., and Trost, W. (2015). Music and emotions: from enchantment to entrainment. *Annals of the New York Academy of Sciences* *1337*, 212–222.
- Wälti, M.J., Woolley, D.G., and Wenderoth, N. (2019). Reinstating verbal memories with virtual contexts: Myth or reality? *PLoS One* *14*, e0214540.
- Wan, C.Y., and Schlaug, G. (2010). Neural pathways for language in autism: the potential for music-based treatments. *Future Neurology* *5*, 797–805.
- Wang, B. (2013). Facial expression influences recognition memory for faces: Robust enhancement effect of fearful expression. *Memory* *21*, 301–314.
- Wang, S., and Agius, M. (2018). The neuroscience of music; a review and summary. *Psychiatr Danub* *30*, 588–594.
- Wang, S., Peterson, D.J., Gatenby, J.C., Li, W., Grabowski, T.J., and Madhyastha, T.M. (2017). Evaluation of Field Map and Nonlinear Registration Methods for Correction of Susceptibility Artifacts in Diffusion MRI. *Frontiers in Neuroinformatics* *11*, 17.
- Warrenburg, L.A. (2020). Choosing the Right TuneA Review of Music Stimuli Used in Emotion Research. *Music Perception* *37*, 240–258.
- Watanabe, K., Masaoka, Y., Kawamura, M., Yoshida, M., Koiwa, N., Yoshikawa, A., Kubota, S., Ida, M., Ono, K., and Izumizaki, M. (2018). Left Posterior Orbitofrontal Cortex Is Associated With Odor-Induced Autobiographical Memory: An fMRI Study. *Frontiers in Psychology* *9*, 687.
- Watanabe, T., Yagishita, S., and Kikyo, H. (2008). Memory of music: Roles of right hippocampus and left inferior frontal gyrus. *NeuroImage* *39*, 483–491.
- Weiss, M.W., Trehub, S.E., and Schellenberg, E.G. (2012). Something in the Way She Sings: Enhanced Memory for Vocal Melodies. *Psychological Science* *23*, 1074–1078.
- White, D., Dunn, J.D., Schmid, A.C., and Kemp, R.I. (2015). Error Rates in Users of Automatic Face Recognition Software. *PLOS ONE* *10*, e0139827.

- Willander, J., and Larsson, M. (2006). Smell your way back to childhood: Autobiographical odor memory. *Psychonomic Bulletin & Review* 13, 240–244.
- Willander, J., and Larsson, M. (2007). Olfaction and emotion: The case of autobiographical memory. *Memory & Cognition* 35, 1659–1663.
- Willander, J., Sikström, S., and Karlsson, K. (2015). Multimodal retrieval of autobiographical memories: sensory information contributes differently to the recollection of events. *Front. Psychol.* 6.
- Williamson, V.J., Baddeley, A.D., and Hitch, G.J. (2010). Musicians' and nonmusicians' short-term memory for verbal and musical sequences: Comparing phonological similarity and pitch proximity. *Memory & Cognition* 38, 163–175.
- Wilmer, J.B., Germine, L., Chabris, C.F., Chatterjee, G., Williams, M., Loken, E., Nakayama, K., and Duchaine, B. (2010). Human face recognition ability is specific and highly heritable. *PNAS* 107, 5238–5241.
- Wilson, D.A., and Stevenson, R.J. (2003). The fundamental role of memory in olfactory perception. *Trends in Neurosciences* 26, 243–247.
- Wilson, D.A., and Sullivan, R.M. (2011). Cortical Processing of Odor Objects. *Neuron* 72, 506–519.
- Wright, D.B., and Sladden, B. (2003). An own gender bias and the importance of hair in face recognition. *Acta Psychologica* 114, 101–114.
- Yardley, L., McDermott, L., Pisarski, S., Duchaine, B., and Nakayama, K. (2008). Psychosocial consequences of developmental prosopagnosia: A problem of recognition. *Journal of Psychosomatic Research* 65, 445–451.
- Yeshurun, Y., and Sobel, N. (2010). An Odor is Not Worth a Thousand Words: From Multidimensional Odors to Unidimensional Odor Objects. *Annual Review of Psychology* 61, 219–241.
- Yeshurun, Y., Lapid, H., Dudai, Y., and Sobel, N. (2009). The Privileged Brain Representation of First Olfactory Associations. *Current Biology* 19, 1869–1874.
- Yin, R.K. (1969). Looking at Upside-down Faces. *Journal of Experimental Psychology* 81, 141–145.
- Yonelinas, A. (1994). Receiver-operating characteristics in recognition memory: evidence for a dual-process model. *Journal of Experimental Psychology: Learning, Memory, and Cognition*.
- Yonelinas, A.P., and Jacoby, L.L. (2012). The process-dissociation approach two decades later: Convergence, boundary conditions, and new directions. *Mem Cogn* 40, 663–680.
- Yonelinas, A.P., Aly, M., Wang, W.-C., and Koen, J.D. (2010). Recollection and Familiarity: Examining Controversial Assumptions and New Directions. *Hippocampus* 20, 1178–1194.
- Young, A.W., and Burton, A.M. (2018a). Are We Face Experts? *Trends in Cognitive Sciences* 22, 100–110.
- Young, A.W., and Burton, A.M. (2018b). The Limits of Expertise in Face Recognition: Response to Sunday and Gauthier's "Face Expertise for Unfamiliar Faces: A Commentary on Young and Burton's 'Are We Face Experts?'" 1, 6.

Young, A.W., and Burton, A.M. (2018c). What We See in Unfamiliar Faces: A Response to Rossion. *Trends in Cognitive Sciences* 22, 472–473.

Yovel, G., and Belin, P. (2013). A unified coding strategy for processing faces and voices. *Trends in Cognitive Sciences* 17, 263–271.

Zarzo, M. (2008). Psychologic Dimensions in the Perception of Everyday Odors: Pleasantness and Edibility. *Journal of Sensory Studies* 23, 354–376.

Zatorre, R.J., and Salimpoor, V.N. (2013). From perception to pleasure: Music and its neural substrates. *PNAS* 110, 10430–10437.

Zhou, G., Lane, G., Cooper, S.L., Kahnt, T., and Zelano, C. (2019). Characterizing functional pathways of the human olfactory system. *ELife* 8, e47177.

(2018). *Handbook of Research Methods in Human Memory* (Routledge).