



Bringing very long term memories back to life

Christelle Larzabal

► To cite this version:

Christelle Larzabal. Bringing very long term memories back to life. Human health and pathology. Université Paul Sabatier - Toulouse III, 2017. English. NNT : 2017TOU30179 . tel-01989938

HAL Id: tel-01989938

<https://theses.hal.science/tel-01989938>

Submitted on 22 Jan 2019

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.



Université
de Toulouse

THÈSE

En vue de l'obtention du

DOCTORAT DE L'UNIVERSITÉ DE TOULOUSE

Délivré par : *l'Université Toulouse III - Paul Sabatier*

Spécialité : *Neurosciences*

Présentée et soutenue le : *07/07/2017* par :

Christelle LARZABAL

BRINGING VERY LONG-TERM MEMORIES BACK TO LIFE

RÉVEILLER NOS MÉMOIRES LES PLUS ANCIENNES

JURY

Rodrigo QUIROGA	University of Leicester	Rapporteur
Aude OLIVA	Massachusetts Institute of Technology	Rapporteur
Laurence TACONNAT	Université de Tours	Examineur
Lluís FUENTEMILLA	Universitat de Barcelona	Examineur
Simon THORPE	Université Toulouse III, Paul Sabatier	Directeur de thèse

Ecole doctorale : *ED CLESCO – Spécialité : Neurosciences*

Unité de recherche : Centre de Recherche Cerveau et Cognition (CerCo)

Directeur de thèse : Simon Thorpe

Bringing very Long-Term Memories back to Life, Christelle Larzabal

ABSTRACT

Memory plays tricks on us and fails us even for recent events. So for the retrieval of sensory information that we have not experienced for decades, surely the memories are gone without a trace, but can we be sure of this? What if the information had been there all along ready to be explicitly retrieved through recall or familiarity for example? So far, experimental evidence is lacking. The purpose of my thesis was to shed some light on the retrieval of these inactive memories. To tackle this problem I developed my research around two main axes: For the first part which is theoretical, I suggest that a) reactivations are not necessary to maintain very long-term memories if the memories were sufficiently repeated at first; b) it might be possible to retrieve explicitly remote inactive memories using specific cues; in the second part which is experimental, I show that a) within specific conditions participants are able to retrieve explicitly very long-term memories that were left inactive for decades and b) such memory traces could be found on the Electroencephalography (EEG) signals. Surprisingly, when conditions are met, remote memories that were thought to be lost can again elicit recall or familiarity. With these results, a question remains: How do neurons store such information?

Keywords: Long-Term memory; Sensory information; Explicit memory; Inactive Memory; Recall; Familiarity; Electroencephalography;

RESUMÉ

On pense souvent que notre mémoire nous joue des tours et nous fait défaut. Ces oublis concernent généralement des souvenirs qui ne sont pourtant pas si vieux. Alors quand il s'agit de se rappeler d'informations sensorielles auxquelles nous n'avons pas repensé depuis plusieurs dizaines d'années, pas de doute, il ne reste plus rien... Mais en est-on vraiment sûr? Et si l'information était toujours là, prête à se manifester de manière explicite -comme sous forme de rappel ou de sentiment de familiarité- pour peu que les conditions le lui permettent? A ce jour, très peu d'expériences ont été menées sur ce sujet. L'objectif de ma thèse a donc été d'apporter des premiers éléments de réponse concernant la récupération de mémoires laissées inactives. Ma recherche s'est orientée selon deux axes principaux : une première partie théorique où je montre a) qu'il n'est pas nécessaire de réactiver une trace mnésique pour la maintenir à très long terme si elle a été suffisamment répétée et b) que la récupération explicite de cette mémoire serait possible grâce à la présentation d'informations qui cibleraient au mieux le souvenir ; une seconde partie expérimentale où je montre a) que des participants sont capables de récupérer de manière explicite des informations laissées en dormance pendant au moins une dizaine d'années et b) que cette trace mnésique pourrait être détectée sur des tracés d'Electroencéphalographie (EEG). Ainsi, et aussi étonnant que cela puisse paraître, dans des conditions favorables à leur réactivation, d'anciennes traces mnésiques que l'on croyait disparues, peuvent de nouveau surgir sous la manifestation de rappel ou de sentiment de familiarité par exemple. Ces résultats soulèvent des questions majeures concernant le stockage neuronal de cette information.

Mots-clés : Mémoire à long-terme ; Information sensorielle ; Mémoire explicite ; Mémoire inactive ; Rappel ; Familiarité ; Electroencéphalographie

RESUMÉ SUBSTANTIEL

Il est communément admis que pour conserver nos souvenirs il vaut mieux y repenser régulièrement. En effet, la réactivation d'une information permet le maintien, voire le renforcement de la trace mnésique sous-jacente. Cependant, dans certains cas, l'information sensorielle peut être perçue suffisamment de fois pour être stockée en mémoire à long terme puis ne plus être réactivée. Est-il alors possible de conserver la trace de ces mémoires qui n'ont plus été réutilisées ? Ou bien sont-elles tout simplement effacées au bout d'un certain nombre d'années ? Récemment, il a été montré que même des images vues à trois reprises et pendant quelques secondes pouvaient encore influencer le comportement de participants presque 20 ans plus tard (Mitchell, 2006). Il semblerait donc que nos traces mnésiques soient toujours disponibles même lorsque celles-ci ne sont pas réactivées. Cependant et comme le fait remarquer Mitchell dans son article, la persistance de ces traces mnésiques a été mise en évidence dans une tâche de récupération implicite où les participants ne savaient pas qu'ils devraient essayer de récupérer en mémoire les images vues auparavant.

Dans ma thèse, j'ai voulu aller plus loin en testant dans des tâches de mémoire explicite s'il serait possible que des participants puissent récupérer des informations sensorielles à très long-terme et laissées dormantes. Dans ces tâches les participants étaient donc poussés à récupérer une information sensorielle bien précise.

Pour ce faire, j'ai orienté ma thèse selon deux axes principaux : un axe théorique et un axe expérimental.

L'axe théorique est composé des chapitres 1 et 2 de ma thèse. Le chapitre 1 qui s'appuie sur une revue de littérature soumise à publication, a pour objectif de montrer que les processus mnésiques d'encodage et de récupération seraient le résultat d'inférences statistiques de type Bayésien s'inscrivant en continuité directe avec nos mécanismes perceptuels. Ces inférences seraient calculées en continu et permettraient l'accumulation d'informations spécifiques à l'origine de la récupération d'un souvenir dans des tâches expérimentales bien définies. Ce modèle rendrait ainsi compte de l'ensemble des phénomènes de récupération mnésique.

Dans le chapitre 2, je développe ce modèle dans le cas de la récupération d'information à très long terme. Je montre notamment qu'il serait possible de récupérer explicitement des mémoires laissées dormantes pendant des dizaines d'années. Pour ce faire, il faudrait arriver à

cibler l'information à l'aide d'indices bien spécifiques qui seraient utilisés dans des tâches les plus favorables à la récupération explicite.

Le second axe, l'axe expérimental est la partie la plus largement développée dans ma thèse et est constituée des chapitres 3, 4, 6 et 7. Les chapitres 3 et 4 correspondent à deux expériences où j'ai testé la récupération explicite d'informations sensorielles laissées dormantes pendant une dizaine d'années au moins. Les données présentées dans ces deux chapitres ont fait l'objet d'articles qui ont été acceptés (chapitre 3) ou soumis à publication (chapitre 4). Dans le chapitre 3, des participants déclarant posséder une télévision durant les années 1960-1970 ($N = 34$) ont été testés sur leur capacité à se souvenir de programmes télévisés ($n = 50$) diffusés durant ces années-là et non rediffusés depuis. Pour chaque programme télévisé, un court extrait du générique était proposé. En moyenne, les participants ont jugé que sept à huit de ces extraits audiovisuels (15%) leur étaient familiers. Parmi ces réponses, deux participants ont été capables de rappeler spontanément le titre de deux de ces programmes télévisés. Lorsqu'un rappel spontané n'était pas possible, les participants étaient invités à identifier le titre parmi quatre propositions. Au total 50 réponses correctes ont été obtenues. Ces réponses correctes étaient associées d'un sentiment de familiarité, d'une certaine confiance dans la réponse que les participants avaient donné, et d'informations contextuelles spécifiques au moment de la diffusion de l'épisode télévisé.

Ces premières données nous permettent de confirmer ce que l'on avait prédit théoriquement, à savoir qu'il est possible de réveiller des informations sensorielles laissées à l'état de dormance pendant plusieurs décennies, suscitant leur rappel ou la récupération explicite d'informations bien spécifiques.

Dans le chapitre 4, nous allons encore plus loin en montrant que la récupération explicite d'informations sensorielles vues en moyenne trois fois et pendant seulement quelques secondes est toujours possible dix ans plus tard. En effet 24 participants qui avaient vus des images très simples ($n = 144$) il y a environ 12 ans ont été capables d'identifier de nouveau ces images parmi un choix de deux propositions avec des performances meilleures que celles du niveau chance. Cet effet s'est vu en moyenne au travers des participants et a été particulièrement visible chez huit d'entre eux.

Les chapitres 6 et 7 correspondent à des expériences où j'ai analysé le signal cérébral mesuré par électroencéphalographie (EEG) pendant la récupération explicite d'information sensorielle à long terme. Pour des raisons pratiques, ces expériences ont été menées sur une échelle de temps bien plus courte que les deux expériences comportementales développées ci-

dessus (cinq mois maximum). Ces expériences ont été menées de manière très similaire : elles comportaient plusieurs sessions d'apprentissages d'extraits audiovisuels très courts et ensuite une session de rappel où les participants étaient invités à juger si l'image qui leur était proposée avait été extraite d'un clip audiovisuel qui leur avait été montré ou non.

Dans le chapitre 6, je regarde tout particulièrement l'effet de l'intervalle de rétention entre le moment de l'apprentissage et le moment de rappel. Je montre ainsi la différence d'activation cérébrale mesurée en EEG chez des participants ($N = 11$) entre des images nouvelles et des images présentes dans des clips ($n = 375$) montrés quelques heures, un jour ou trois semaines auparavant. Dans le chapitre 7 je regarde l'effet du nombre de présentations des clips audiovisuels lors du rappel trois semaines plus tard. Dans ce cas, les clips ($n = 480$) étaient montrés soient une, deux, trois ou cinq fois chez des participants ($N = 12$). Cette expérience s'est poursuivie d'une seconde où la récupération en mémoire de clips audiovisuels a été testée cinq mois plus tard, dans un paradigme où les clips audiovisuels étaient remontrés aux participants.

L'analyse des données récoltées en EEG s'est faite selon des mesures bien spécifiques que je développe dans le chapitre 5 de ma thèse. Je montre ainsi que les techniques d'analyses multivariées de patterns permettent d'identifier des traces mnésiques à long terme.

En conclusion, cette thèse permet de révéler que nombreuses traces mnésiques seraient conservées à l'état latent. En élaborant des tâches expérimentales bien spécifiques, il serait possible de réveiller de manière explicite des informations sensorielles répétées seulement quelques fois. Ces résultats questionnent le substrat biologique de ces mémoires. Serait-il possible de conserver des traces au sein d'un vaste réseau neuronal ? Nous espérons que de nouvelles recherches permettront d'apporter de futurs éléments de réponse.

ACKNOWLEDGEMENTS

Je me permets d'écrire cette section dans mon français natal.

Tout d'abord je tiens à remercier Simon Thorpe qui dès le début m'a donné son entière confiance dans l'élaboration de ce projet de thèse. Je me souviens de son enthousiasme –et de sa chemise Hawaïenne !- lors de notre premier entretien visiophonique. J'avais postulé à une seule offre de thèse, celle que Simon proposait, et dès les premières minutes de notre conversation, je savais que c'était avec lui que je voulais faire ma thèse. En faisant partie de son équipe j'ai eu la chance d'avoir tout le support matériel et humain qui m'a permis de conduire l'ensemble de mes recherches dans des conditions plus qu'optimales. Malgré ces quatre années de thèse, je dois avouer que je connais encore très mal le milieu de la recherche mais je pense difficilement me tromper en disant que Simon est un directeur de thèse atypique. Toujours plein d'enthousiasme, Simon est prêt à tester n'importe quelle idée aussi farfelue qu'elle puisse paraître. Et des idées ce n'est pas ce qui lui manque ! C'est donc dans ce bouillonnement intellectuel que ma thèse s'est peu à peu construite.

La deuxième personne que je tiens à remercier est Nadège Macé. Nadège avait fait sa thèse avec Simon quelques années auparavant et était la manager de l'équipe pendant la première moitié de ma thèse. C'est grâce à Nadège que j'ai pu créer la structure de mon projet de thèse, me lancer dans les expériences et les analyser. Nadège était là à tous les niveaux de mon travail et permettait de faire le lien entre Simon et ma thèse. Mais bien au-delà de ces apports professionnels, Nadège est pour moi une personne de confiance, qui m'a écouté et soutenu dans les moments difficiles où la thèse n'était pour moi plus qu'une préoccupation mineure. Suite aux décisions ministérielles, Nadège a été forcée de quitter l'équipe, un départ qui nous a tous affecté.

Après le départ de Nadège, c'est Karine Bouyer qui est arrivée et qui a été la manager pendant la deuxième partie de ma thèse. Je tiens à remercier Karine pour toute son aide au niveau administratif et pour avoir toujours pris le temps de faire un point concernant l'avancée de mon travail. Tellement pratique que son bureau soit juste à côté du mien et de celui de Simon !

J'ai eu la chance que mon bureau soit également en face de celui d'Emmanuel Barbeau. Un merci tout particulier à Emmanuel et à Eve Tramoni d'avoir voulu collaborer sur un projet qu'ils avaient en tête depuis plusieurs années. L'encadrement d'Emmanuel sur ce projet, m'a permis d'acquérir une plus grande rigueur scientifique lors de chaque étape de ce travail.

Ce bureau justement, j'ai pu le partager pendant ces quatre années avec Jaya, ma compère de thèse et récemment docteur, Sophie, qui m'a assisté dans les passations de toutes les expériences que j'ai menées au CerCo, Evelyne qui était en post-doc pendant une petite année, Danaé, qui a commencé sa thèse depuis un an et Méyi, en stage de M2. Un grand merci à vous pour la super ambiance qu'il y a eu dans ce bureau avec beaucoup, beaucoup de rires ! Je remercie également Elodie et Martin, pour toute la bonne humeur dans le bureau à côté.

Je tiens également à remercier Maroufa pour ses gâteaux, ses petits plats mijotés et surtout nos petites conversations qui me donnaient toujours le sourire à la fin d'une longue journée de travail.

Plus généralement, je tiens à remercier le M4 team et l'ensemble des personnes qui sont au CerCo ou ont été au CerCo. C'est grâce à vous tous, que la bonne ambiance au labo était possible. Un merci tout particulier à Pier Zanone dont les conseils ont été précieux pour les quelques dernières retouches apportées à mon manuscrit.

Il serait bien trop réducteur cependant de se limiter à la sphère CerCo.

Je tiens à remercier tous les participants qui ont bien voulu consacrer un peu de leur temps pour mes expériences. Tous ces résultats ne seraient pas là sans eux !

Je remercie également l'ensemble des membres du jury pour avoir accepté d'évaluer mon travail. Merci à Rodrigo Quiroga, à Aude Oliva, à Laurence Taconnat et à Lluís Fuentemilla.

Un grand MERCI aussi à Chris Moulin qui m'a donné la chance de collaborer avec lui. Merci, de m'avoir accueillie à Grenoble pendant un mois, et merci d'avoir consacré autant de temps sur notre projet. Je tiens à le dire, ma thèse aurait été bien moins aboutie sans tout cet apport intellectuel.

Je veux aussi en profiter pour remercier Jean-Rémi King, Simon van Gaal et Martijn Wooke, qui ont permis de tracer mon chemin jusqu'à cette thèse et Simone pour m'avoir ouvert un tout nouvel horizon après cette thèse.

Et pour finir un énorme merci à tous mes amis : Cécilia et Leila en particulier, la Coloc Sansou ainsi que le groupe des Dimitri de Pau et mes amis du Pays Basque.

MILESKER

Merci à ma famille,

Merci Callista.

Table of Contents

ABSTRACT	3
RESUMÉ.....	5
RESUMÉ SUBSTANTIEL.....	7
ACKNOWLEDGEMENTS	11
Table of Contents.....	15
Table of Figures.....	21
GENERAL INTRODUCTION	25
1. The essence of the words	25
1.1 Brain	25
1.2 Information: In-formation	26
1.3 A few musical notes on memory	27
2. Studying very long-term memories	28
2.1 Explicit memories	29
2.2 A brief literature overview	30
2.3 Overall framework.....	33
2.4 Presentation of the chapters.....	34
• PART 1 - THEORETICAL FRAMEWORK -	37
CHAPTER 1. Unifying memory and perception into the same framework... 39	
Abstract	39
Chapter presentation	40
1. Introduction	41
2. Memory encoding and retrieval: A dynamic Bayesian framework.....	43
2.1 Memory encoding: from perceptual processing to the creation of a distinct percept	44
2.2 Memory retrieval: a dynamic Bayesian account	47
2.3 The hippocampus: a repository of memory with high prior values	55

3. Evidence for a dynamic Bayesian model	56
3.1 Timing of early processes	58
3.2 Timing of higher level processes	59
4. Clinical application	61
5. Conclusion.....	64

CHAPTER 2. From recent and active memories to remote and inactive ones

Abstract	67
Presentation chapter	68
1. Introduction	70
2. Memory consolidation	71
2.1 Synaptic consolidation	71
2.2 Repetition: the key for neuronal selectivity and long-term storage	72
2.3 System-level consolidation	75
3. Post-consolidation reactivations, prior and accessibility	75
3.1 Memory activity in a Bayesian perspective	76
3.2 Updating the prior value	77
4. Silent neurons for dormant memories: experimental evidence	79
4.1 From STDP to dynamic Bayesian inferences	79
4.2 Inactive neurons for inactive memories	82
5. Conclusion.....	84
• PART 2 – EMPIRICAL FRAMEWORK -	85
Presentation of the behavioral experiments.....	87

CHAPTER 3. Waking up buried memories of old TV programs.....

Abstract	89
Presentation chapter	90
1. Introduction	91

2.	Material and Methods	93
2.1	Participants	93
2.2	Stimuli	93
2.3	Task	94
3.	Results	95
3.1	Participants' performance for the famous TV programs	95
3.2	Older participants' performance for the test TV programs	97
3.3	Influencing factors	105
4.	Discussion	106
5.	Conclusion	109
6.	Acknowledgements	109
7.	Supplementary information	110
CHAPTER 4. Long-term memory and familiarity after 12 years		115
	Abstract	115
	Presentation chapter	116
1.	Introduction	119
2.	Method	120
2.1	Participants	120
2.2	Initial test: the DMS-48	120
2.3	Current tasks	122
3.	Results	126
3.1	Naïve control participants	126
3.2	Test participants' performance	126
4.	Discussion	132
5.	Acknowledgment	134
	Discussion of the behavioral results	135
	Presentation of the imaging experiments	137

CHAPTER 5. Detecting the retrieval of long-term memories on a EEG signal

1. Introduction	139
2. EEG compared to other neuroimaging techniques	140
2.1 EEG	140
2.2 Comparison	140
3. Description of the techniques used	142
3.1 ERPs analyses	142
3.2 Time-Frequency analyses.....	143
3.3 MVPA for decoding subtle effects	145
4. Conclusion.....	148

CHAPTER 6. Decoding long-term memories of short audiovisual clips..... 149

Abstract	149
Chapter presentation	150
1. Introduction	152
2. Materials and methods	153
2.1 Participants.....	153
2.2 Stimuli.....	153
2.3 Procedure	153
2.4 EEG acquisition and analysis	155
3. Results	157
3.1 Behavioral results	157
3.2 EEG analysis	161
4. Discussion	165
5. Conclusion.....	167
6. Acknowledgements.....	168
7. Supplementary information.....	168

CHAPTER 7. Effect of repetition on explicit long-term memory: an EEG study	171
Chapter presentation	171
1. Introduction	172
2. Material and Methods	173
2.1 Participants	173
2.2 Stimuli	173
2.3 Task	174
2.4 EEG acquisition and analysis	175
3. Results	177
3.1 Behavioral results	177
3.2 EEG analysis	180
4. Discussion	182
5. Material and Methods	183
5.1 Participants	183
5.2 Stimuli	183
5.3 Task	184
6. Results	185
6.1 Behavioral results	185
7. Discussion	186
8. Supplementary information	187
GENERAL DISCUSSION	189
1. Wrap-up	189
1.1 Overall framework	189
1.2 Substantial progress	189
2. Theoretical considerations	190
2.1 Identification, familiarity, novelty	190
2.2 Consciousness and memory	191
2.3 Concepts and neuronal selectivity	192
2.4 Forgetting	194
2.5 False memories and confabulations	198

2.6	The role of affects	200
3.	Future directions	200
3.1	Limits	200
3.2	Linking behavior and decoding analyses	201
3.3	From Bayesian to integrated information.....	201
4.	Conclusion.....	203
4.1	A personal view to my thesis	203
4.2	Can one see the shape of our memories?	203
OPENING. Beyond matter, from Alan Turing to Cymatics		205
1.	Standing waves	205
2.	Cymatics.....	206
3.	Harmonics in the brain.....	207
FINANCIAL SUPPORT		209
REFERENCES.....		211

Table of Figures

Figure I.1 Representations of the cortical structures	26
Figure I.2 Schematic representations of the hippocampus and the rhinal cortices	29
Figure I.3 An illustration of the forgetting curves	32
Figure 1.1 From perception to memory.....	42
Figure 1.2 Memory retrieval: a decision based on the evidence accumulated over time towards different hypotheses/classes.	50
Figure 1.3 Decisions based on the evidence accumulation model.	57
Figure 1.4 Two-dimensional representation of the distribution of the neuronal responses on a Yes/No judgement and a on a forced-choice task.	64
Figure 2.1 Illustration of the Sharpening model.	73
Figure 2.2 Retrieval from inactive to active memories.....	78
Figure 2.3 Activations of neurons during the presentation of a specific stimulus.....	81
Figure 3.1 Experimental design.	95
Table 3.1 Collection of 50 reports for the test videos judged familiar by the older participants.....	103
Figure 3.2 Performance in the 4-FC for the 50 test TV programs.....	104
Table 3.S1 Older participants' detailed information: gender, age, year of TV acquisition and daily number of hours of watching TV at the time of the recall.....	110
Table 3.S2 Details of famous and test TV programs: broadcast period, number of episodes, duration of episodes and number of seasons.....	113
Figure 4.1 Examples of pairs of stimuli used in the DMS-48 for the three conditions "Unique", "Paired" and "Abstract".	121
Figure 4.2 The current experimental procedure consisted in two explicit memory tests: a forced-choice (task 1) and a Yes/No recognition (task 2).	124
Figure 4.3 Test participants' performance on the current experiment.	127
Figure 4.4 Participants' performance on the forced-choice and Yes/No tasks according to the number of times stimuli were seen in the DMS-48 (2002-2008): less than three times (< 3) or 3 times and more (≥ 3).	128
Figure 4.5 Participants performance over the confidence level in the forced-choice task.	130

Figure 5.1 Spatiotemporal scale of neural signals for commonly used neuroimaging techniques.	139
Figure 5.2 Lay-out of a 64-electrode cap using the BIOSEMI system.....	140
Figure 6.1 Experimental design.	155
Figure 6.2 Behavioral performance: HITs responses on average across participants...159	
Figure 6.3 Differential EEG activity between snapshots seen in the 3 learning sessions (HOURS, 1D and 3W before the recall) and NEW snapshots.	163
Figure 6.4 Early ERPs analysis on average across 11 participants for snapshots seen HOURS or 1D before and compared to New pictures.	164
Figure 6.5 Decoding analysis for the comparisons between the 3 learning session (Hours, 1D, 3W) and New pictures on average across participants.....	165
Figure 6.S1 Differential EEG activity between REMEMBER or KNOW snapshots and NEW snapshots.....	169
Figure 6.S2 Decoding analysis for the comparisons between Remember or Know responses and New pictures on average across participants.....	170
Figure 7.1 Experimental design.	175
Figure 7.2 Participants recall performance for videos seen 1, 2, 3 or 5 times 3 weeks before.	179
Figure 7.3 Differential EEG activity between snapshots seen ONE, TWO, THREE or FIVE TIMES, three weeks ago and NEW snapshots.	181
Figure 7.4 Decoding analyses for the comparisons between the four numbers of exposures (ONE, TWO, THREE, FIVE) and NEW pictures on average across participants.....	182
Figure 7.5 Experimental design.	184
Figure 7.6 Retrieval performance over the number of exposures for audiovisual clips seen five months ago.	186
Figure 7.S1 Differential EEG activity between REMEMBER or KNOW snapshots and NEW snapshots.....	187
Figure 7.S2 Decoding analyses for the comparisons between the four numbers of exposures (ONE, TWO, THREE, FIVE) and NEW pictures on average across the five participants with the best decoding performance.	187
Figure O.1 Exemplar of Chladni’s figures (from (Chladni, 1787)).....	206

“We watch an ant make his laborious way across a wind- and wave-molded beach. He moves ahead, angles to the right to ease his climb up a steep dune let, detours around a pebble, stops for a moment to exchange information with a compatriot. Thus he makes his weaving, halting way back to his home. So as not to anthropomorphize about his purposes, I sketch the path on a piece of paper. It is a sequence of irregular, angular segments not quite a random walk, for it has an underlying sense of direction, of aiming toward a goal. I show the unlabeled sketch to a friend. Whose path is it? An expert skier, perhaps, slaloming down a steep and somewhat rocky slope. Or a sloop, beating upwind in a channel dotted with islands or shoals.

Perhaps it is a path in a more abstract space: the course of search of a student seeking the proof of a theorem in geometry.

Whoever made the path, and in whatever space, why is it not straight; why does it not aim directly from its starting point to its goal? In the case of the ant (and for that matter the others) we know the answer. He has a general sense of where home lies, but he cannot foresee all the obstacles between. He must adapt his course repeatedly to the difficulties he encounters and often detour uncrossable barriers. His horizons are very close, so that he deals with each obstacle as he comes to it; he probes for ways around or over it, without much thought for future obstacles. It is easy to trap him into deep detours. Viewed as a geometric figure, the ant's path is irregular, complex, hard to describe. But its complexity is really a complexity in the surface of the beach, not a complexity in the ant.

On that same beach another small creature with a home at the same place as the ant might well follow a very similar path.”

Herbert A. Simon, The Sciences of the Artificial, 1996

GENERAL INTRODUCTION

1. The essence of the words

1.1 Brain

Brain: an organ of soft nervous tissue contained in the skull of vertebrates, functioning as the coordinating center of sensation and intellectual and nervous activity (Oxford definition).

This definition goes back to the biological structure of the brain from which high level functions emerge. At the end of its development, the brain is composed of several structures: the brain stem: with the midbrain, the pons and the medulla, the diencephalon with the thalamus and hypothalamus, the cerebellum and the cerebrum. The cortex, the most external layer of gray matter of the cerebrum is divided into neocortical and allocortical structures, the latter including the hippocampus and the rhinal cortices.

From an evolutionary perspective, the cerebral cortex can be seen as a tissue that was constrained to develop inside rigid boundaries delimited by the skull. This expansion was particularly important in primates and especially in humans. To overcome this space limitation, the tissue adapted its growth by folding up into gyri (ridges) and sulci (depressions). To some extent, gyri are used to delimit cortical areas that have a structural and a functional specificity. As an example, the inferior temporal gyrus which corresponds approximately to the Brodmann area 20 is particularly involved in recognition memory. Cortical areas are included in four main lobes: frontal, parietal, temporal, occipital (figure I.1). But as already observed by Ramon y Cajal at the beginning of the 20th century, the relative uniformity of the six-layered structure of the neocortex demonstrates its structural unity.

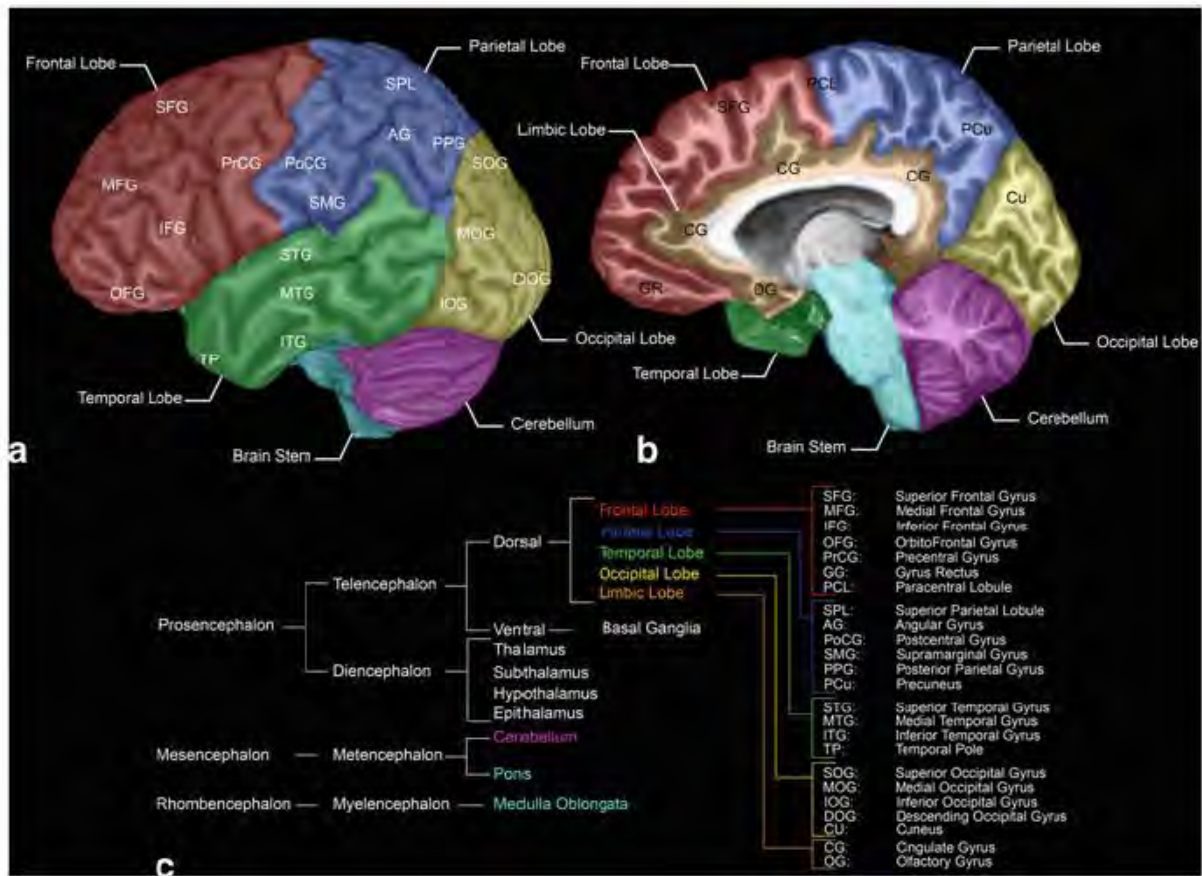


Figure I.1 Representations of the cortical structures

Lateral view (A), medial view (B) and brain ontogeny (C) (from (Forstmann et al., 2015))

1.2 Information: In-formation

The vibrations of a djembe after being hit by a musician

The severe damage on the trunk of the tree following the storm





My body is still shivering after you kissed me on the neck

A tissue was in a certain condition but is now in a different state due to physical perturbations coming from the environment. This transition shows the integration of new information into the system through the reshaping or the deformation of the tissue.

Like for every organic tissue, the brain is exposed to environmental constraints and is constantly deformed by the action of external or internal agents. Deformation allows the integration of new inputs by the tissue and through its content creates the IN-FORMATION necessary to perceive the world. According to Bohm, “Information is a difference of *form* that makes a difference of *content*, i.e. meaning” (Bohm, 1989). To explain it simply, Bohm gives

the example of letters on a printed page: the difference of meaning is held in the difference in the forms of the letters.

Therefore, any type of information would be revealed through the interaction between two entities. If we consider human experiences, information would be generated from the interaction between an external object (external input) and the nervous tissue or between an internal input and the nervous tissue. While it seems obvious that any external object has a specific form, the same would apply for internal inputs. Indeed, internal inputs, commonly called thoughts or mental images would originate from the direct abstraction of objects present in the outside world. Such new forms might be more subtle than the ones we experience in the outside world, yet it does not restrict the potential of information they can convey. Again, to explain it simply, writing could be a good example.

Nowadays, it is hard to see any relationship between letters and their corresponding physical objects. However such correspondence does exist. Let's take the example of the two first letters of the Latin alphabet: "A/a" and "B/b" which correspond to alpha and beta in the Greek alphabet. Alpha derives from the Phoenician alphabet "aleph" , which means head of ox, and this comes directly from the Egyptian hieroglyph . Beta is equivalent to the "bet" letter  in the Phoenician alphabet which means house and derives from the Egyptian hieroglyph .

It is interesting to see that every Egyptian hieroglyph corresponds to a symbol that is a direct abstraction of an object taken from the outside world. In a way, a living ox or a real house would be the external objects and the Egyptian hieroglyphs or the letters A and B, their respective mental images.

The form of these "mental images" can correspond to ideas (logograms) or to sounds (phonograms) that combined together provide specific information and can have as much influence as real objects.

One would easily understand why such writing was considered as sacred at the time: the close relationship between the symbols and their outside attributes gave them sacred powers which would be activated when read.

1.3 A few musical notes on memory

Information is in the shape and can be called *quale* (Balduzzi and Tononi, 2009) in the field of neuroscience. A shape is informative because it can be characterized by two invariants: its frequency range of resonance also called spectrum and the whole set of possible points in that

frequency range. A point in that shape can be seen as a chord and the whole set of possible chords as the musical scale of the shape.

Every time some information is integrated in our brain specific shapes might be created. In this perpetual movement which never ends and which does not really stabilize, the brain tries to create stable representations based on the information provided by this continuous ballet. This choreography of shapes would not have the same substance as the forms we are used to experiencing in our daily life and could not be reached by our sense of vision. Instead, a more subtle approach would be needed. The major scale with its seven notes (do-re-mi-fa-sol-la-si / C-D-E-F-G-A-B) is a beautiful example of such idea by corresponding to the vibration of the quantum sphere. As such, any representation we build in mind at any time might correspond to the vibration of a specific shape that has emerged from the structural organization of assemblies of neurons. Some of these experiences can be frivolous and would not have any impact on the nervous tissue while others might have a stronger influence. In the latter case, if the information leads to the modification of the structural organization of the tissue it results in the encoding of the input. The creation of this memory trace can be seen as an imprint in the brain. As a natural extension, memory could be viewed as the preservation of those shapes generated by specific neuronal organizations and remembering, literally: being “again member”, the reinstatement of those shapes into conscious access. Accordingly, human knowledge would be held in subtle shapes that could be used in specific domains. From this view, one would better understand why in the Greek mythology, Mnemosyne, the goddess of Memory was also the mother of arts, science and literature embodied in the nine muses.

Importantly, the maintenance of these memory traces over time would vary and only specific constraints would enable the creation of very long-term memories and their potential retrieval. As I will show in the next section, little is known about our ability to retrieve very long-term memories explicitly. In this yet to be explored field, my thesis will try to provide a better insight.

2. Studying very long-term memories

Classically, (very) long-term memories are split into two main categories: explicit memories that can be verbalized (declarative) and implicit memories which are reflected in individual's behavior (nondeclarative) (Squire and Zola-morgan, 1991). In my thesis, I will focus on explicit memories that rely on the Medial Temporal lobe unlike implicit memories.

2.1 Explicit memories

2.1.1 The Medial Temporal Lobe

As its name suggests, the Medial Temporal Lobe is part of the temporal lobe which is one of the four cortical regions (the green region in the figure I.1). It includes allocortical areas including the hippocampus that is part of the archicortex and the rhinal cortices that are part of the periarthicortex, a transitional zone between the archicortex and the neocortex (figure I.2). The hippocampus is composed of the dentate gyrus, four CA fields and the subiculum. The rhinal cortices include the entorhinal and perirhinal cortices. Interestingly, the Medial Temporal Lobe is also made of neocortex with the parahippocampal structure located posteriorly to the rhinal cortices.

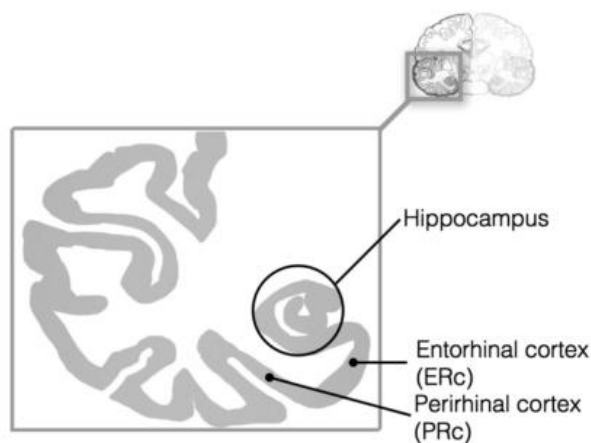


Figure I.2 Schematic representations of the hippocampus and the rhinal cortices

From (Moulin et al., 2013))

The Medial Temporal lobe is a critical region from which phenomenological experiences of explicit retrieval can emerge. It can lead to familiarity if one can categorize an item as previously encountered (it is an “old” item) or to recollection if the categorization as an “old” item is associated with contextual details (Mandler, 1980).

2.1.2 Semantic, episodic, autobiographic

Explicit memories can be semantic if they reflect general knowledge such as retrieving a name, or episodic if a unique specific event can be retrieved within its spatial and temporal context such as the first day in college (Tulving, 1985a). When semantic and episodic memories are related to the self, they form autobiographical memories that reflect the past of

every individual's life. Therefore, all autobiographical memories are not necessarily episodic (Levine et al., 2002; Piolino et al., 2002; Tulving et al., 1988).

2.1.3 Assessing explicit memories

Explicit memories are potentially verbalized by the participants. Therefore they have to be tested in a task where the participants are voluntarily engaged in the retrieval of specific information. This way, participants' performance is a direct result of a retrieval process that they initiated. Several paradigms have been designed to test explicit retrieval and can be split into two main categories: recall tasks and recognition tasks. In recall tasks, participants have to retrieve a specific piece of information that is not present in their environment (free recall) or only partially (cued-recall). In recognition tasks the information is presented to the participants -either alone in a Yes/No task or among several propositions in a forced-choice task- and participants are asked to identify which item was previously encountered.

As we see, the amount of information necessary to perform these tasks significantly varies and not being able to retrieve a name in a free recall task would not necessarily mean not being able to identify it from two instances. This will be discussed at length in chapter 1.

2.2 A brief literature overview

The memory literature is huge but when it comes to very long-term or remote memory -that is for periods of a few months to decay- less data exist. This is mainly due to practical reasons: while it is easy to test the retrieval of sensory information the same day as it was presented, it is more difficult to make sure that all conditions will be met again to test the same participants several years later. Despite this strong constraint, some studies were able to test the retrieval of very long-term memories.

2.2.1 Lifetime memories

One possible way for explicitly testing participants' very long-term memories, is to ask them retrieve information acquired in their real life. Using this first method, researchers could potentially test lifetime memories for a wide range of "ecological" material related to semantic or episodic information such as the memory of old classmates (Bahrick et al., 1975), of information learnt in school (Bahrick, 1984) or of specific autobiographical events (Rathbone et al., 2008; Rubin and Schulkind, 1997; Viard et al., 2007). These studies showed that explicit information acquired during young adulthood could be maintained through the individual's lifetime.

Interestingly, data from amnesic patients revealed that very long-term memories could be robust to damage in the Medial Temporal lobe (Bayley et al., 2006, 2003; Reed and Squire, 1998) although this is not very clear for very long-term episodic information (for a review see: (Frankland and Bontempi, 2005)). This ambiguity led to the contemporary debate: standard model versus Multiple Trace Theory Model (MTT). According to the standard model, the Medial Temporal Lobe (MTL) and the hippocampus in particular would be involved in supporting declarative memories (semantic or episodic) that were recently formed but over time the memories would become independent of the hippocampus and would only rely on cortical structures (Squire and Alvarez, 1995). For the MTT, this would be true only for semantic information and remote episodic memories would still rely on the hippocampus (Nadel and Moscovitch, 1997).

2.2.2 Forgetting curves

When possible, the amount of information retrieved in a given task can be assessed for different retention intervals. The data can be used to draw forgetting curves with on the x-axis of the graph the time elapsed since the acquisition phase and on the y-axis the percentage of retrieval in the specific task (see figure I.3 for an example). This graphical representation provides valuable information on participants' performance over long time delays even if they do not necessarily encompass the whole lifespan of an individual (for example: (Conway et al., 1991; Squire, 1989)). In particular, forgetting curves showed that retrieval performance quickly decreases after the learning phase but then stabilizes for several decades. It is hard to say which mathematical law follows any of the forgetting curves but the power function of time seems to fit well with the data (Anderson and Schooler, 1991; Rubin and Wenzel, 1996; Wixted and Ebbesen, 1991).

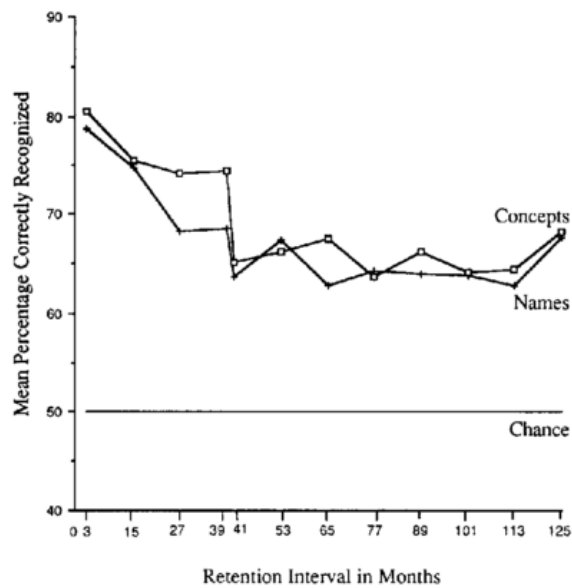


Figure I.3 *An illustration of the forgetting curves*

The percentage of correctly recognized concepts and names is plotted for various retention intervals. From (Conway et al., 1991).

2.2.3 Long-term memories: the where under the scope of fMRI studies

The stabilization of long-term memories comes with a reorganization of its underlying neuronal network. This system-level consolidation is a slow process taking at least several weeks to be completed. As we mentioned, two main theories: the standard model and the Multiple Trace Theory have been proposed to account for this reorganization. This debate, not yet closed, provided a great enthusiasm inside the whole neuroscience community. Many investigations were carried out to get a better understanding of the role of the hippocampus over long retention intervals. The spatial resolution provided by the fMRI technique was particularly helpful and provided evidence toward the Multiple Trace Theory (Harand et al., 2012; Maguire and Frith, 2003; Maguire, 2001). As we see, the role of the hippocampus on long-term retrieval focused most of the attention. This would explain the poor amount of data concerning the temporal dynamics during long-term retrieval assessed via EEG (Electroencephalography) or MEG (Magnetoencephalography). Indeed, only a couple of studies examined participants' EEG activity during the retrieval of items seen the previous day (Curran and Friedman, 2004; Jaeger et al., 2010; Wolk et al., 2006) and only one study went as far as recording the EEG activity for pictures seen four weeks before (Tsivilis et al., 2015).

2.2.4 Remaining questions

This brief survey shows that the explicit retrieval of long-term memories that have been acquired several decades ago is possible. However, one main limitation comes from the fact that the experimenter has little control on the acquisition parameters during new learning. Among them, two parameters are difficult to assess: the strict number of times the information was presented and the fact that the information could have been encountered again outside the original acquisition period. One exception might be the television test that was designed to evaluate remote memories of one-season TV programs used in several studies by Squire et al. (Squire and Fox, 1980; Squire and Slater, 1975; Squire, 1989).

Several questions remain unanswered: How long can information that has not been revisited since the initial acquisition be maintained? Does this duration depend on the number of exposures during the acquisition phase? How do brain processes unfold over time during the retrieval of remote memories? These questions are at the core of my thesis.

2.3 Overall framework

My thesis would not have been possible without the scientific environment in which it took place. After completing my Master's degree, Simon gave me the opportunity to join his team for a new and exciting research project: the M4 project. "M4" stands for Memory Mechanisms in Man and Machine but "Man" should now be changed to "Mammals" with Danaé Rémon's work on dogs. The substantial financial support that Simon got from the European Research Council for five years (FP7/2007-2013/ERC grant agreement n° 323711) allowed him to hire several PhD students, post-docs, managers, and a research assistant to create the M4 team. The idea was to gather people from different backgrounds –psychology, neuroscience or computation-, to tackle the issue of the formation and the maintenance of very long-term memories. Simon used his strong expertise in visual perception and in fast visual recognition in particular (Kirchner and Thorpe, 2006; Thorpe et al., 1996; Wu et al., 2010) together with advanced computational developments (Masquelier and Thorpe, 2007; Masquelier et al., 2009) to think memory from a complete different perspective. Ten claims are at the core of his radical proposal that the M4-team was invited to investigate:

- 1/ Humans can recognize visual and auditory stimuli that they have not experienced for decades;
- 2/ Recognition after very long delays is possible without ever reactivating the memory trace in the intervening period;

- 3/ These very long-term memories require an initial memorization phase, during which memory strength increases roughly linearly with the number of presentations;
- 4/ A few tens of presentations can be enough to form a memory that can last a lifetime;
- 5/ Attention-related oscillatory brain activity can help store memories efficiently and rapidly;
- 6/ Storing such very long-term memories involves the creation of highly selective “Grandmother Cells” that only fire if the original training stimulus is experienced again;
- 7/ The neocortex contains large numbers of totally silent cells (“Neocortical Dark Matter”) that constitute the long-term memory store;
- 8/ Grandmother Cells can be produced using simple spiking neuronal networks models with Spike-Time-Dependent-Plasticity (STDP) and competitive inhibitory lateral connections;
- 9/ This selectivity only requires binary synaptic weights that are either “on” or “off”, greatly simplifying the problem of maintaining the memory over long periods;
- 10/ Artificial systems using memristor-like devices can implement the same principles, allowing the development of powerful new processing architectures that could replace conventional computing hardware.

2.4 Presentation of the chapters

With no doubt, this manifesto largely inspired the direction and the content of my thesis. With Simon and Nadège Macé, the M4-manager during the first half of my thesis, we agreed that I would carry experiments concerning the retrieval of remote memories. Like an archeologist my goal was to find evidence for very long-term memories of sensory stimuli that were thought to be lost. And for that, I had to respect one major constraint: stimuli should not have been presented again for years after the acquisition period. In this adventure I had the support from the National Institute of Audiovisual in Toulouse (chapter 3), and from Emmanuel Barbeau and Eve Tramoni who gave me access to a list of participants that we could test in a well-controlled design (chapter 4). I am very grateful to Sophie Muratot who helped me collecting all this precious data.

In such a pre-defined framework I could have easily fallen in the trap of becoming a mere “inspector” that would validate (or reject) one of the claims above. Lucky me, Simon gave me the opportunity to express myself in this project. My main motivation was to understand the underlying processes at the basis of our subjective experiences during memory retrieval. Especially my interest lied in its dynamic component. I decided to carry EEG experiments to study how processes unfold in time during long-term retrieval (chapter 5). After the conclusive results of the first EEG experiment (chapter 6) I decided to follow-up and to study

the impact of repetition for long-term retrieval (chapter 7). This project was successfully handled by Méyi Dulème who presented the results in her Master's thesis. Gathering experimental evidence was a first step in understanding the dynamic of retrieval, another step was to propose a theoretical framework that could explain it. In this enterprise I had all the support from Chris Moulin. It was an extraordinary chance for me that Chris had to evaluate my work at the mid-term of my thesis. Following that, he gave me the opportunity to work on a literature review about the dynamics of retrieval (chapter 1). This helped me to prepare the ground for the topic of very long term retrieval (chapter 2).

2.4.1 Peer-reviewed articles

Some chapters of this thesis correspond to articles that were submitted for peer-review in scientific journals. For easier reading, I slightly adapted them in my thesis.

Unifying memory and perception into the same framework: a dynamic Bayesian approach

Christelle Larzabal, Simon J. Thorpe, Chris J.A Moulin

Review article submitted

Waking up buried memories of old TV programs

Christelle Larzabal, Nadège Bacon-Macé, Sophie Muratot, Simon J. Thorpe

Research article accepted in Frontiers in Behavioral Neuroscience (2017)

Long term memory and familiarity after 12 years

Christelle Larzabal, Eve Tramoni, Sophie Muratot, Simon J. Thorpe, Emmanuel J. Barbeau

Research article submitted

The landscape is portrayed, the road is opened. I invite you now to follow me on this path.



- **PART 1 - THEORETICAL FRAMEWORK -**



CHAPTER 1. Unifying memory and perception into the same framework

Abstract

Traditionally, memory is split into separate systems and subsystems according to their function. Although perception can also be similarly broken down into separate subsystems, many neuroscientific views of perception tend to see processing as a meaningful whole, with information flowing through a series of stages from low level processing in V1 to conceptual categorization in the medial temporal lobe. In this article we draw upon such perceptual theories to propose that encoding and retrieval of information is based on the ‘accumulation of evidence’, arising directly from perceptual mechanisms. As such, we see memory as a single system, but like perception, it involves varying levels of processing based on the computational nature of the task in hand, such as subjective awareness, discrimination, categorization, and retrieval of context. Episodic memory and recollection lie at the top of this information processing hierarchy. We review evidence from experimental data in support of this hypothesis and show how it might account for clinical pathologies such as Alzheimer’s disease.

Chapter presentation

This first chapter reflects in many aspects all the knowledge I acquired during my Master's degree and that I had time to fully process to integrate it in the field of memory. I started with subliminal and bi-stable perception with Jean-Rémi King in Paris, before studying the role of intuition in decision-making with Martijn Wokke in Amsterdam. I have to tell that my 6-month internship with Jean-Rémi was an invaluable experience. With him I really started to understand and do cognitive neuroscience. He introduced me to many notions, either theoretical like Bayesian inferences, or technical like decoding analyses applied to MEG data. Coming from the field of biology all of this was new to me but so exciting! The following year I had the chance to go to Amsterdam for a 6-month internship. I remember the first paper Martijn told me to read: The neural basis of decision Making (Gold and Shadlen, 2007). It was quite tough and so abstract at first! I did not give up reading and finally I got the essence of it. A decision was the result of Bayesian inferences computed over time. While decision-making is usually associated with motor actions such as going to a specific direction, it is a very generic term that also holds for perception or memory. Indeed the task that our participants had to perform was like an implicit memory task. It is with this background that I started my PhD with for main focus the retrieval of very long-term memories.

I quickly realized that I was not the only one who started from perception and ended up with memory. It was also the case of Simon who just hired me! I knew that we would share the same vision: the idea of a continuum between perception and memory. However since the beginning Simon was reluctant to the idea of Bayesian inferences. I remember when I discussed it with Chris during my mid-thesis committee. He was more interested and suggested to write a review on this topic. We dedicated a lot of time for this project. I even went to visit Chris in Grenoble for a month to work on this project full time. I thank Karine, the lab manager, for making my stay easy in Grenoble.

The discussion went on for months with Simon and finally we ended up with a review article: "Unifying memory and perception into the same framework: a dynamic Bayesian approach" that we submitted for peer-review. It has the merit of being a personal view rooted on theoretical notions about memory and tries to make the bridge between two main cognitive fields: perception and memory. Note that for ease of reading some sections of the review article were skipped and a few changes were provided.

1. Introduction

Encoding and retrieving from human memory are complex activities involving many dispersed brain regions. For instance, the frontal lobe is often described as being involved in the control and monitoring of memory (Goldman-Rakic, 1987; Miller, 2000), the parietal lobe is recruited in recognition/episodic memory and decision making (Adolphs et al., 1996; Cabeza et al., 2008; Platt and Glimcher, 1999; Shadlen and Newsome, 1996), and the occipital lobe is implicated in the generation of visual imagery and in perceptual priming experiments (Kosslyn et al., 2001; Roland and Gulyas, 1994; Wiggs and Martin, 1998). However, the medial temporal lobe (MTL) - including the hippocampus and the entorhinal, perirhinal and parahippocampal cortices - is usually considered to be the critical area for memory function. In particular, the hippocampus is known to be a crucial brain region; if it is removed or damaged, memory function is severely impaired with an inability to either store new information (anterograde amnesia) or retrieve distant memories (retrograde amnesia). The hippocampus is thus seen as playing a critical role in memory (Burgess et al., 2001; Marr, 1970; Rolls, 1996; Squire, 1992). Contemporary views of hippocampal function see it as particularly crucial for episodic memory, the ability to retrieve contextual specifics from memory. Some authors consider that it is only involved in the initial formation of representations which later rely on neocortical sites (the standard model: e.g. (Squire et al., 1984)) whereas others see it as a central processor that remains involved in the binding of separate sources of information together into a coherent whole (Olsen et al., 2012; Ryan et al., 2000; Yonelinas, 2013). Here, we will argue that these views concerning hippocampal activation and binding can be linked with theories of perception.

Interestingly, the hippocampus is located at the end of the visual ventral stream and receives input from multiple sensory modalities (Suzuki and Amaral, 1994). The hippocampus is also a strategic center that plays a key role in the integration of sensory inputs (Squire et al., 2004). Based on this neuroanatomical structure, our hypothesis is that memory lies on a continuum with perceptual processes, and that the encoding of information emerges from the routine treatment of perceptual input. After having been processed in subcortical regions such as the lateral geniculate nucleus (LGN) in the thalamus, sensory inputs have to pass through successive cortical areas for the information to be integrated. For visual stimuli (see Figure 1.1), the input follows the visual ventral stream by entering the primary visual area V1, where the neurons fire for low-level properties of stimuli (e.g. bars, edges). Higher order attributes are then encoded through the next regions of the visual cortex (V2, V3) until the Lateral

Occipital Complex (LOC) in humans or the Inferior Temporal cortex (IT) in macaques. It is in the LOC that we find the most specific or selective neurons in the pathway, such as face-selective neurons. The hippocampus, at the end of this perceptual stream, is thought to be involved in binding of sensory information to other stored representations, and it is here that multi-modal selectively firing neurons have been found (such as the ‘Halle Berry’ cell which fires for her face, name and for the character she plays in Catwoman (Quiroga et al., 2005). By receiving monosynaptic hippocampal projections (Thierry et al., 2000), direct communication between the prefrontal cortex and the hippocampus is possible and occurs preferentially during the first stages of the formation of the memory trace: encoding and consolidation (Benchenane et al., 2010). Recent decoding techniques used on combined MEG and fMRI data reveal a dynamic continuum of activation associated with an increase of categorization from V1 to IT (Cichy et al., 2014) that might continue until the MTL and PFC. This strongly suggests that early M/EEG effects cannot differentiate different types of stimulus, but as activation spreads throughout the brain, it allows the entanglement of such representations (DiCarlo and Cox, 2007).

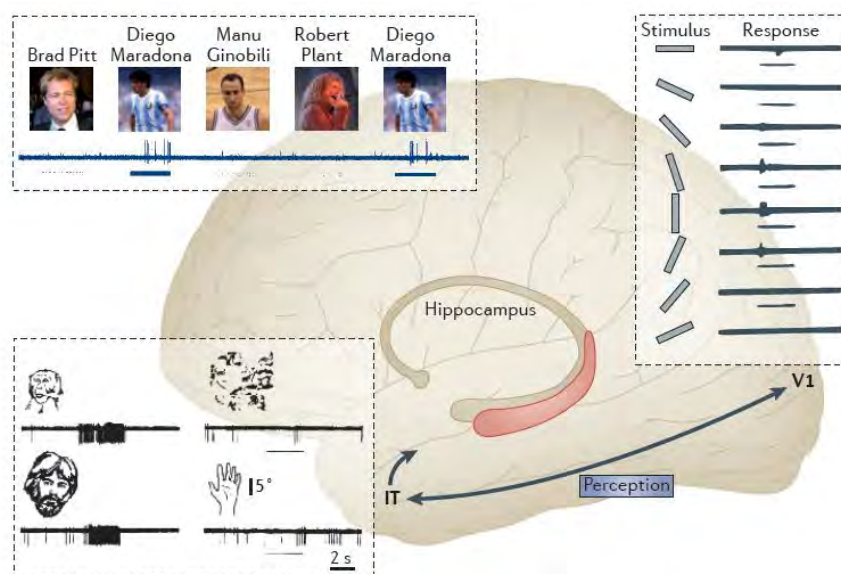


Figure 1.1 From perception to memory.

In the visual ventral stream the processing of visual stimuli starts in the primary visual area V1, where the neurons fire for low-level properties of stimuli such as for a specific bar orientation (a vertical bar in the figure). The neuronal selectivity increases through the next regions and reaches its maximum at the end of the visual ventral stream: in the Inferior Temporal cortex (IT) in macaques or in the Lateral Occipital Complex (LOC) in humans. At this stage, neurons encode higher order attributes such as specific body parts (faces in the

figure) that can be bound to other representations in the medial temporal lobe and in the hippocampus in particular, where neurons fire for specific concepts (the footballer Maradona in the figure) (from (Quiroga, 2012))

We make the novel suggestion that this evolution of cells firing according to the differentiation of increasingly specific and complex representations (which can be thought of as a perceptual process) could ultimately lead to further differentiation, creating a distinct percept that contains all information associated with that percept. We propose here that this process of differentiation is akin to the perceptual process of evidence accumulation over time, such that memory is based on the outputs of a perceptual system making ever fine-grained evaluations of items in perception.

We suggest that memory encoding and retrieval are dynamic processes that unfold over time and that directly emerge from perceptual mechanisms. We will show that dynamic Bayesian inferences can explain both memory encoding and retrieval by relying mainly on the MTL. In particular we will show how the interaction between retrieval tasks (e.g. free recall or forced-choice tasks) and retrieval assessments (e.g. Remember/Know/New judgement) lead to different retrieval modes (naming, feeling of knowing, familiarity...) that emerge from the same evidence accumulation model. We will review evidence from behavioral and time-resolved experiments that support this model. We strongly believe that this model can also be extended to clinical cases and as an example, we will also illustrate how this evidence of accumulation model might account for Alzheimer's disease.

2. Memory encoding and retrieval: A dynamic Bayesian framework

First proposed by von Helmholtz in 1866 (Helmholtz, 1947) perceptual integration is now often considered as an inference problem and can be explained in Bayesian terms (Kersten et al., 2004; J-R. King and Dehaene, 2014; Ko and Lau, 2012). Given the noisy information coming from our senses at a time t and prior knowledge acquired through experience, the brain's role is to decide what the most probable sensory input is. This process of reaching a suitable decision based on priors is referred to as the accumulation of evidence. The implication is that over time, Bayesian processes drive the acquisition of information that enables a final perceptual decision. Here we propose a natural continuum of this idea, and propose that memory encoding and retrieval can be similarly modelled as a Bayesian process. Note that in itself, the idea of priors in perceptual decision making is, in essence, about memory, since priors are in some way merely an agglomeration of activations based on

previous experience. Even the very idea of identifying a stimulus rests upon the idea that something in the environment is matched with an internal representation, which is again, based on memory.

2.1 Memory encoding: from perceptual processing to the creation of a distinct percept

2.1.1 A dynamic Bayesian approach

During the encoding of a stimulus, a cascade of brain regions are activated which are reflected by the firing of neuronal assemblies associated with an alteration of their synaptic weights. As with perceptual decision-making we suggest that the encoding of a stimulus is supported by a succession of brain processes performing Bayesian statistics that decide at their own level which representation is the most probable given the incoming stimulus. Depending on the environmental circumstances, perceptual decisions will tend to focus on specific features and will reduce the possibilities into several hypotheses; part of this decision making will involve long term memory—such as decisions about the classification of objects (e.g. living or non-living), their identification (e.g. cat or dog) and simple judgements of familiarity.

In our view, this diminution of possibilities over time leads to a reduced and clustered multidimensional space, where the brain's role is to find the most probable hypothesis H_i given noisy sensory evidence E coming from the selected features. This inference problem is expressed in Bayesian terms as finding $P(H_i|E)$, the posterior probability. The posterior is defined as being proportional to two parameters that can be known, the likelihood function and the prior:

$$P(H_i|E) \sim P(E|H_i)P(H_i)$$

With $P(E|H_i)$: the *likelihood* function of obtaining evidence from the sensory input;

$P(H_i)$: the probability of occurrence of the stimulus H_i called *prior*;

The posterior probability is calculated for all different hypotheses: H_1 to H_x : $P(H_1|E)$, $P(H_2|E)$... $P(H_x|E)$ with the purpose of extracting the one with the maximal *a posteriori* probability.

Perceptual decisions and decisions in general are not conducted on a single maximum *a posteriori* value. Instead, to be sure that posterior value has some consistency the same calculation is performed over a certain period of time. The idea is to accumulate evidence in support of a hypothesis (Gold and Shadlen, 2007). Mathematically, this is done by

multiplying the logarithm of the likelihood ratios (log LR, called also weight of evidence) obtained at each time i (from 1 to n values). The prior can thus be considered as being the log LR at time 0.

For a binary decision this is expressed as:

$$\text{Log LR}_{12} \equiv \log \frac{P(e_1, e_2, \dots, e_n | h_1)}{P(e_1, e_2, \dots, e_n | h_2)} = \sum_{i=1}^n \log \frac{P(e_i | h_1)}{P(e_i | h_2)}$$

This means that at each time i , a decision variable ($DV = \log LR$) is calculated from the evidence obtained up to that point. Decisions can be made if a certain criterion is reached or if the cost is too high to continue the process (Gold and Shadlen, 2007). This mathematical approach might explain the differences observed in terms of subjective report according to the criterion level used. Indeed, explicit access would only emerge if the DV reaches a subjective threshold that corresponds to a high level criterion. If the DV is too low, but still differs from random noise, implicit behaviors can be observed (Dehaene, 2008). In addition, every DV comes with an estimate of its level of uncertainty that might be the basis of metacognitive judgments. Metacognition commonly described as ‘knowing what we know’ (or what we do not know) does not only apply to explicit processes and can be considered as a second order judgement performed on unconscious or conscious processes (Charles et al., 2013; Spehn and Reder, 2000). A metacognitive estimate of certainty can be translated in terms of confidence levels or error detection and is used to initiate and co-ordinate efficient memory behaviors (Moulin and Souchay, 2015).

2.1.2 *The creation of a conceptual representation*

During memory encoding, low-level regions such as V1 and V2 for the visual ventral stream are first activated by an incoming stimulus. Depending on the stimulus properties and according to the task in which participants are involved, brain activity can spread out to higher cortical structures such as in IT before reaching the prefrontal cortex and the MTL structures. It has been shown in the monkey that the proportion of cells with activity that directly corresponds to subjective perception increases along the visual ventral stream: from 20% in V1 to 90% in IT (Logothetis, 1998). This strong link has also been seen in studies of epileptic patients (Kreiman et al., 2002; Quiroga et al., 2008b) where most of the neurons in the MTL were activated according to the patient's subjective reports, the MTL getting direct input from IT. The MTL has been implicated in the last stages of visual processing as shown by deficits in visual spatial cognition in patients with Alzheimer's disease (Butter et al., 1996;

Possin, 2010). But it is important to point out that the MTL does not seem to take part in fast perceptual decisions. Indeed, perceptual recognition tasks involving differentiating broad categories such as Animal/Non Animal can be achieved in 150 ms (Thorpe et al., 1996) and saccades to animals can be made in as little as 120 ms (Kirchner and Thorpe, 2006), which is far faster than the latencies of MTL neurons (~300 ms).

- The hippocampus as a “timer”

Why does it take so long for the hippocampus to react? In principle, it is just a couple of synapses at most, so conduction delays should not be sufficient. Here we suggest that one of the key roles of MTL structures, and of the hippocampus in particular, could be to keep track of the duration of neuronal activations occurring in IT (Reddy and Thorpe, 2014; Thorpe, 2011), at the end of the visual ventral stream, as well as in other high-level cortical regions such as in LIP, at the end of the visual dorsal stream, or in other multisensory areas. According to this idea, though strong, transient activations of cortical neurons would not be long enough for hippocampal neurons to kick in and only long-lasting activations (of approximately 150-200 ms) would trigger the firing of specific hippocampal neurons. Although speculative, this hypothesis gives a simple explanation of the extreme long latency of hippocampal neurons -about 300 ms - which have been recorded in epileptic patients (Mormann et al., 2008). Therefore, only neuronal representations that remain active for 150 ms will get recorded in the hippocampus and elicit perceptual awareness (Quiroga et al., 2008b). These long-lasting activations would lead to the creation of conceptual representations both in the hippocampus and in the neocortex that could be explicitly recalled. In the case of multimodal stimuli (e.g. audiovisual clips) we suggest that these long-lasting cortical activations are processed simultaneously in a global neuronal workspace (Baars, 1988; Dehaene et al., 1998a) where for example, spatio-temporal information from the dorsal pathway, emotional content from the amygdala and self-awareness about the particular representation encoded in the prefrontal cortex can be bound together to create a distinct percept (Capurro and Quiroga, 2009). This would explain why binding occurs in the presence of perceptual awareness (Crick and Koch, 1990; Engel and Singer, 2001). At around 300 ms, depth P3b potentials have been recorded in several brain regions with the largest located in the hippocampus (Halgren et al., 1998). Scalp and depth P3b are generally presented as ERP components reflecting conscious processes (Silverstein et al., 2015).

We suggest that broadly, this activation can be considered as reflecting the achievement of the encoding process with the creation of a stored-representation accessible to the subject. This

means that later on this stored-representation will have the capacity of being recalled explicitly (verbally, for instance). By contrast, in the case of transient activations, the absence of specific hippocampal activity will not allow the formation of any conceptual representation that could be explicitly recalled. Nevertheless, because information was encoded in low-level structures, changes in perceptual abilities like fluency facilitation (Dehaene et al., 1998b; Marcel, 1983) can emerge, demonstrating implicit retrieval.

As explained above, we consider that the encoding of a stimulus is supported by a cascade of brain processes performing Bayesian statistics to decide at their own level which representation is the most probable given the incoming stimulus. It seems to be an efficient process as demonstrated by the fact that resulting outcomes are usually close to our real-world experience except for some cases like visual illusions (Geisler and Kersten, 2002). In the following section, we argue that memory retrieval shares a lot in common with the way a representation is encoded. More specifically we will use the same dynamic Bayesian framework to explain the different retrieval levels than can emerge from a search into the memory space.

2.2 Memory retrieval: a dynamic Bayesian account

2.2.1 Anatomical considerations

The hippocampus plays a dual role in human memory, and is involved in both the encoding and retrieval of information through two different pathways (Treves and Rolls, 1992). As confirmed by experimental data (Leutgeb et al., 2007) the monosynaptic pathway from the entorhinal cortex to the CA3 region of the hippocampus enables pattern completion, the ability to retrieve past information from incoming inputs. On the other hand, the disynaptic pathway recruiting in that order the entorhinal cortex, the dentate gyrus and the CA3 region, could allow pattern separation and the ability to encode new stimuli.

We will argue that in addition to simply being located in the same cerebral structures, encoding of new memories and their retrieval are processed through the same dynamic Bayesian framework which implies an accumulation of knowledge over time and across judgement types in different experimental paradigms.

2.2.2 *A dynamic Bayesian model for memory retrieval*

We start our argument with reference to recognition memory, where ultimately, a decision about prior occurrence is made based on perceptual information in the environment. Importantly, in terms of the medial temporal lobe memory system, and most particularly in recognition memory, episodic memory judgements can be decomposed into familiarity and recollection (Aggleton and Brown, 1999; Eichenbaum et al., 2010). However, the nature of these two different retrieval modes remains unclear as illustrated by the debate between a dual and single process accounts (Mandler, 2008; Moulin et al., 2013) and reflects the limits of the current memory theories. We think that a novel model might be useful to explain these different levels of memory access.

Although a dynamic Bayesian framework has been widely used to model perceptual integration of sensory inputs over time and consequently fits well with the Bayesian brain hypothesis (Knill and Pouget, 2004), surprisingly, this perspective has rarely been mentioned for understanding memory retrieval. One exception comes from the diffusion decision model developed by Ratcliff in the late 1970s (Ratcliff, 1978) which can be applied to both perception and memory (see (Ratcliff et al., 2016) for an updated version). In this model, evidence for a stimulus is accumulated over time and a decision is made when a certain threshold/criterion is reached. The criterion corresponds to the minimal amount of evidence required to perform above chance level in a given task, and the amount of evidence directly reflects the level of information access. In memory retrieval we have to account for a broad range of processes, including the capacity to process items more fluently the second time we encounter them (implicit memory), but also including free-recall (such as being able to recall Jennifer Aniston's name) and the accessing of previous events (such as being able to recall the last time that we saw Jennifer Aniston). We suggest that these retrieval modes lie on a continuum, representing levels of memory accessibility (Tulving, 1985b). In our model, higher levels of accessibility in memory are accompanied by higher levels of evidence. One strength of our model is to consider participants' subjective responses according to the related experimental task as recently proposed for perceptual judgments (J-R. King and Dehaene, 2014; Ko and Lau, 2012) and this will vary dynamically over the time-course of the decision.

The successful retrieval of information varies according to the amount of environmental support provided by the context. This factor directly determines the amount of effort needed to retrieve any particular information: the self-initiated activity (Craik, 1986; Lindenberger

and Mayr, 2014). High levels of environmental support which require little self-initiated activity are found when the stimulus appears in the same form as in the environment, such as in a recognition task, and the subject merely has to report whether they have seen the stimulus or not. In contrast, free recall tasks involve more self-initiated activity because of the low level of environmental support available. This directly translates in terms of performance where participants have better memory scores in recognition tasks than for free recall (Lindenberger and Mayr, 2014).

We will focus on four paradigms commonly used in the literature: two that are related to recall tasks: free recall and cued-recall, and two relying on recognition processes: the Yes/No task and the forced-choice task. These four paradigms involve, in this order, decreasing amounts of self-initiated activity. Placed in a Bayesian framework, we propose that the amount of self-initiated activity triggered by the retrieval task reflects the number of hypotheses from which the decision is drawn: the less environmental support is provided, the more hypotheses must be considered. The top of the posterior probability for these different hypotheses is represented in Figure 1.2 by colored circles. The placement of the criterion that best separates the different hypotheses into different classes allows the generation of an adapted response. We will argue that two different criteria are enough to explain the variety of retrieval phenomena: one for familiarity and one for ‘knowing’. They are respectively depicted by blue and orange lines in Figure 1.2. Note that knowing here corresponds to semantic knowledge concerning facts about the world (semantic generic-knowledge) or about our personal life (semantic self-knowledge) that can be explicitly expressed through naming, for example. It refers to noetic consciousness in Tulving terms (Tulving, 1985a) and should not be confused with Know responses from a Remember/Know/New paradigm that often reflect familiarity responses (Gardiner and Richardson-Klavehn, 2000; Williams et al., 2013).

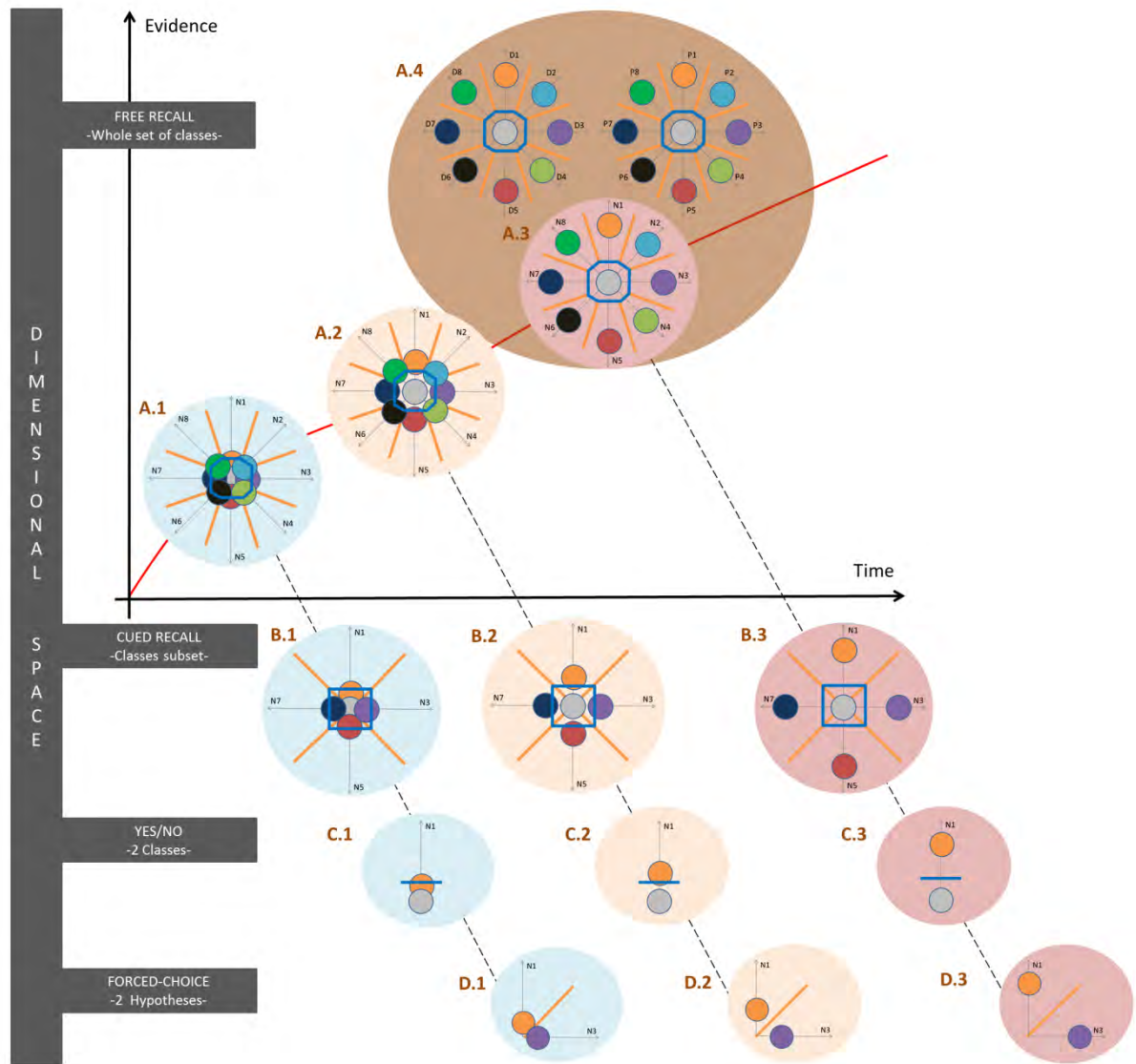


Figure 1.2 Memory retrieval: a decision based on the evidence accumulated over time towards different hypotheses/classes.

The top of the posterior probability calculated for each hypothesis is represented by a colored circle in the dimension of the retrieval decision. Dimensionality is represented in the figure by the number of possible stimuli involved in retrieval. The two potential criteria that are used in the decisional process: “familiarity” and “knowing” are represented by the blue and orange lines respectively. Over time the class distributions get more and more differentiated and facilitate retrieval access. Decision is made on all possible hypotheses for free recall (plot A.1 from A.4) and is restricted to a subset of classes on a cued recall (plot B.1 from B.3). For recognition tasks the number of classes is highly reduced with to 2 classes for a Yes/No task (plot C.1 to C.3) and 2 hypotheses for a forced-choice task (plot D.1 to D.3). Several examples of retrieval access are detailed: Absence of retrieval (A.1), Familiarity (A.2),

Knowing (A.3) and Remembering (A.4) in a free recall task, knowing (B.2 and B.3) in a cued-recall task, familiarity judgement in a Yes/No task (C.2 and C.3) and categorization in a forced-choice task (from D.1 to D.3).

- Free recall

Consider a question such as, *who won the French Open Tennis Championship's women singles title in 2013?* It is a free recall task: a name has to be retrieved without more information. Dealing with that question is not easy because it implies searching in a memory space containing all the possible and accessible hypotheses for the feature “Name” which is huge. As an illustration, in the top of Figure 1.2 are represented the posterior probability for eight hypotheses corresponding to eight different names (colored circles) plus the posterior probability distribution corresponding to new items (grey circle). Retrieval corresponds to a response based on a knowing criterion (orange line) that corresponds here to the ability to report a name. Depending on the amount of evidence accumulated toward a name, correct naming can be performed (diagram A.3, Figure 1.2) or not (diagram A.1 and A.2, Figure 1.2). Indeed, semantic knowledge emerges when all the possible hypotheses are well dissociated from each other.

- Cued recall

Free recall can be facilitated by adding a cue (see (Nairne, 2002) for a contemporary discussion of cue-target relations). A cue is a partial amount of information directly related to the one that has to be retrieved. In our example, it could be the first 3 letters of the name: *Wil*. The cue’s effect will be to reduce the number of possible hypotheses to a subset of classes and thus facilitate retrieval (diagram B.1, B.2 and B.3). Behaviorally, this can be translated in terms of shift from a simple familiarity judgement to the retrieval of a name (diagram B.2 in Figure 1.2).

Information can be added to such an extent that the stimulus presented is not merely a part of the stimulus (a cue) but the stimulus itself. The task no longer consists in trying to recall an item but relies on the recognition of this item as it appears in the environment. Yes/No and the forced-choice recognition tasks are two typical designs to test recognition.

- Yes/No task

In a Yes/No paradigm we have to judge if the item presented on its own was previously encountered or not. In keeping with our example, the stimulus in the Yes/No task might correspond to the name Serena Williams. The task would be to decide whether the fact: ‘Serena Williams was the Roland Garros’ winner in 2013’ is indeed information that we have stored in memory (have encountered before), or is an incorrect, ‘new’ statement. This judgement is performed on two classes of response: “Old” and “New” can merely rely on the familiarity criterion depicted in blue (diagrams C.1, C.2 and C.3, figure 1.2) that applies to a single dimensional feature (one name). At low levels of evidence (diagram C.1, Figure 1.2) the two distributions are both under the familiarity criterion, since not enough evidence has accumulated to distinguish an old item from a new one. If more evidence is accumulated toward the “Old” item, distributions will disperse from the “New” class and correct identification can be based on a familiarity judgement (diagrams C.2 and C.3, Figure 1.2).

- Forced-choice tasks

In a forced-choice task, the number of possibilities is highly reduced: there are only two hypotheses (two possible stimuli), and it is for this reason that it is probably the easiest of all memory decision tasks (e.g. (Bayley and Wixted, 2008; Westerberg et al., 2006). In our example, we might pick one of the two proposed names: Serena Williams and Victoria Azarenka. We again plot the posterior probability reduced to a 2D representational space (bottom plot). On these diagrams the orange diagonal represents the knowing criterion for discriminating the two stimuli such that at low levels of evidence, we can differentiate between two distributions in the representational space, although the distributions overlap (diagram D.1, Figure 1.2) and were not different under the familiarity criterion in the Yes/No task (diagram C.1, Figure 1.2). As more evidence is accumulated, distributions will be more spread apart.

2.2.3 *Evaluation of responses: Another continuum*

Interestingly, for a given memory task, the amount of environmental support is also modulated according to the way retrieval is evaluated. Retrieval measurement is not necessary as categorical as right/wrong and can involve a graduation of states that are reflected in the amount of evidence accumulated. This is particularly visible in recall tasks where a large amount of evidence has to be accumulated. As an example, participants involved in a free recall task perform better when asked to report correctly the Feeling of Knowing for a name than when required to report the actual name (Koriat and Levy-Sadot, 2001). Moreover, in the

above four tasks, we have outlined the lowest level of evidence required to perform certain memory evaluations. It is clear that where excess information exists this will have a bearing on retrieval processes (such as being able to recall the correct answer in any case when we are presented with the material in a recognition task).

In free recall, not being able to give the correct name does not necessarily mean not being able to retrieve any other information. Retrieval judgements can take into account other parameters. The first one relies on a familiarity criterion (blue line) which is less conservative than a knowing criterion in a free recall task because it is only based on two classes: “Old” and “New”. Therefore, if enough evidence is computed for the “Old” class, familiarity toward the response can emerge. In the diagram A.2 (Figure 1.2) the eight hypotheses corresponding to the “Old” class are sufficiently spread out from the “New” class to report a familiarity judgement.

The second parameter involves a knowing criterion that is based on other features related to the response. Indeed, a search for a name into a main feature dimension can be associated with a search for other items from secondary feature dimensions (cf. The Source Monitoring Framework; e.g. (Johnson et al., 1993)), that involve either personal detail (self-knowledge): *I remember watching it in a pub in Paris. I was with other friends of mine* or more generic facts: *It was a short match. A very sunny day.* Remembering an event can therefore be considered as the retrieval of items along several feature dimensions that are closely linked. This search of complementary items (e.g. dates and places) follows the same evidence accumulation model and is made in parallel with the accumulation of evidence performed on the main feature dimension. Depending on the strength of the maximum a posteriori computed in these secondary dimensions, retrieval of specifics can emerge. If retrieval of related other features occurs before knowing the response it might give rise to a Feeling Of Knowing for the item (FOK) (Hart, 1965) (not represented in the Figure 1.2); if it occurs after identifying the response or close in time, it might trigger the remembering of the experience (diagram A.4, Figure 1.2). According to Tulving’s definition, when some of these items refer to self-knowledge, remembering might give rise to autonoetic consciousness and the overall experience being part of episodic memory (Tulving, 1985a).

2.2.4 *First and Second-order judgements*

According to this dynamic Bayesian model we have shown that novelty, familiarity, FOK, semantic knowledge and remember responses can be considered as first-order judgements that are only based on 2 criteria: familiarity and knowing. These 2 criteria might be supported by different brain structures. For example, the entorhinal cortex in the MTL seems to be a perfect candidate to support the familiarity criterion because of its crucial role in the process of familiarity judgements (Aggleton and Brown, 1999; Brandt et al., 2016; de Vanssay-Maigne et al., 2011). On the other hand, the (pre)frontal cortex and the hippocampus might be good candidates for the knowing criterion given their importance in episodic and semantic memory. Interestingly, this idea of considering episodic and semantic memory through the same criterion is in agreement with recent neurobiological findings that showed that episodic and semantic retrieval might involve the same brain network and therefore might reflect variation along a continuum of processing (Burianova et al., 2010; Ryan et al., 2008). More generally frontal structures are proposed to play a role in strategic and conscious memory control (e.g (Moscovitch, 1992; Wheeler et al., 1997)) and these control aspects relate to our higher order ‘knowing’ criteria. In keeping with this idea, Souchay et al. (Souchay et al., 2000) showed that a decline of frontal functions during aging is associated with a decrease in the ability to generate episodic memory FOK. This is consistent with our prediction that FOK might emerge from processes based on a knowing criterion supported by frontal and hippocampal structures. Neuroimaging support for this frontal-hippocampal involvement in metacognition comes from Genon et al. (Genon et al., 2016).

Our model is also close to the standard Signal Detection Theory used to explain the difference between Remember and Know responses, here referred as remember and familiarity judgements, in classical Remember/Know/New paradigms (Dunn, 2004; Wixted, 2009). However, it deviates from it in such a way that differences are not explained by confidence levels (higher confidence levels for Remember compared to Know responses) but in terms of levels of information accessibility. Indeed and as already stressed for perceptual judgements, during memory retrieval it is important to dissociate first-order judgements (such as novelty, familiarity, semantic/episodic memory) that are based on the class with the highest a posteriori for a given task, from second-order confidence judgements that compute to what extent the decision is correct or not (J-R. King and Dehaene, 2014). The further the class is from the decision criterion, the more confident will be the response. Crucially, confidence

judgements should always be considered according to the criterion used because low confidence on remember responses might not equate with high confidence of familiarity for example. Although first and second-order judgements are interrelated, they are not interchangeable (Rajaram et al., 2002), a notion that an experimenter should keep in mind when explaining the different response modalities to participants (Williams et al., 2013).

The main strength of this dynamic representation is to show that the different types of memory retrieval are due to a difference in terms of information accessibility that are part of a continuum going from categorization to a full recollection.

The prior probability distribution of a specific memory refers to what extent we expect to use/access a memory again. This value is high for newly encoded events and decreases over time, unless the memory is revisited. We suggest that one of the key roles of the hippocampus would be to keep track of the memories with high prior values in such a way that it could be considered as a dictionary of recently experienced events (Thorpe, 2011). In the next section physiological evidence is discussed in support of this hypothesis.

2.3 The hippocampus: a repository of memory with high prior values

Intracranial recordings in epileptic patients showed that neurons in the hippocampus are very selective to recently seen stimuli like celebrities (Quiroga et al., 2005), TV series on air such as the Simpsons (Gelbard-Sagiv et al., 2008) or members of the patients' family. This selectivity seems to develop very quickly as shown by the recording of neurons that were selectively responding to previously unknown members of the experimental team within just a few days (Viskontas et al., 2009) and might also disappear quickly if the stimuli are not experienced anymore leaving space for more recent events. We can imagine that if we could record a few months later the same hippocampal neuron that was found selective to an experimenter during a 2-week recording, we would notice a change in its selectivity. The same would apply for the famous Jennifer Aniston neuron if we do not rehearse the memory of the actress.

Although speculative, we can suppose that the prior value could correspond to the number of cells devoted to a particular item in the hippocampus. Over time and in the absence of rehearsal, the number of cells dedicated to a stimulus would decrease, and consequently its

prior value. As a result, memories with low prior values in the hippocampus would be overwritten by newly encoded memories. In contrast, every time a stable memory would be revisited, its memory trace would be reactivated and its prior value would be boosted. Therefore, memories that are very often rehearsed would be associated with a strong prior value that makes these memories more distinguishable from other ones and therefore more easily retrievable. Importantly, reactivations do not necessarily result in explicitly retrieving the memory trace. Some memory traces can be only reactivated in an implicit manner. Low activations of neurons are localized in cortical sites and prior value increases but not enough for the hippocampus to take into account that memory. Interestingly, it has been shown that the reactivation of a stable memory can sometimes be associated with its reconsolidation (Dudai, 2012). Reconsolidation is in fact considered as an update of previous learnt information (Alberini, 2005; Dudai and Eisenberg, 2004; McKenzie and Eichenbaum, 2011) and can lead to its distortion (Sara, 2000). Frequently rehearsed experiences might be therefore completely falsified after some time, a critical point of interest when considering eye witness testimony in courtrooms (Lacy and Stark, 2013).

3. Evidence for a dynamic Bayesian model

Before technology gave us sophisticated devices to investigate the brain from the inside, the only way for scientists to get an idea of the series of processes that were involved in a task, was to rely on behavioral responses and in particular participants' reaction times (Donders, 1969). Experiments were conducted by varying specific parameters and reaction times were compared between conditions.

This is a very reliable methodology that is still widely used in many laboratories. Indeed, as opposed to subjective reports that are coarse-grained judgements, reaction times give more subtle details and especially can inform about processes that are not experienced consciously. Cognitive processes are usually made well before the motor action that closes up the behavior. To get closer to the earliest stages and observe very tiny differences between conditions, it can be better to use saccadic responses than more traditional manual responses (Kirchner and Thorpe, 2006). Nowadays, tasks using manual or saccadic responses can be combined with EEG brain recordings. This is a powerful tool because it allows us to see how brain activity unfolds over time with millisecond precision. As we see, reaction times can tell us a lot about the underlying brain processes. According to the diffusion model (Ratcliff, 1978) whereas timing is represented on the x-axis of a graph representing the time-course of a decision, on the y-axis it is the evidence that it is plotted. Evidence can be thought of as the amount of

goal-oriented information that has been accumulated since the beginning of the Bayesian process. In every computational task: problem-solving but also perception, memory or future prediction, the brain is involved in a decision process. We can then directly infer the amount of evidence that has been reached at a specific time and for a given task. Applied to memory encoding and retrieval, it allows them to be considered as a continuum (Figure 1.3). In this section we will review experimental evidence that has been collected until now and that support our evidence accumulation model hypothesis described in the previous sections.

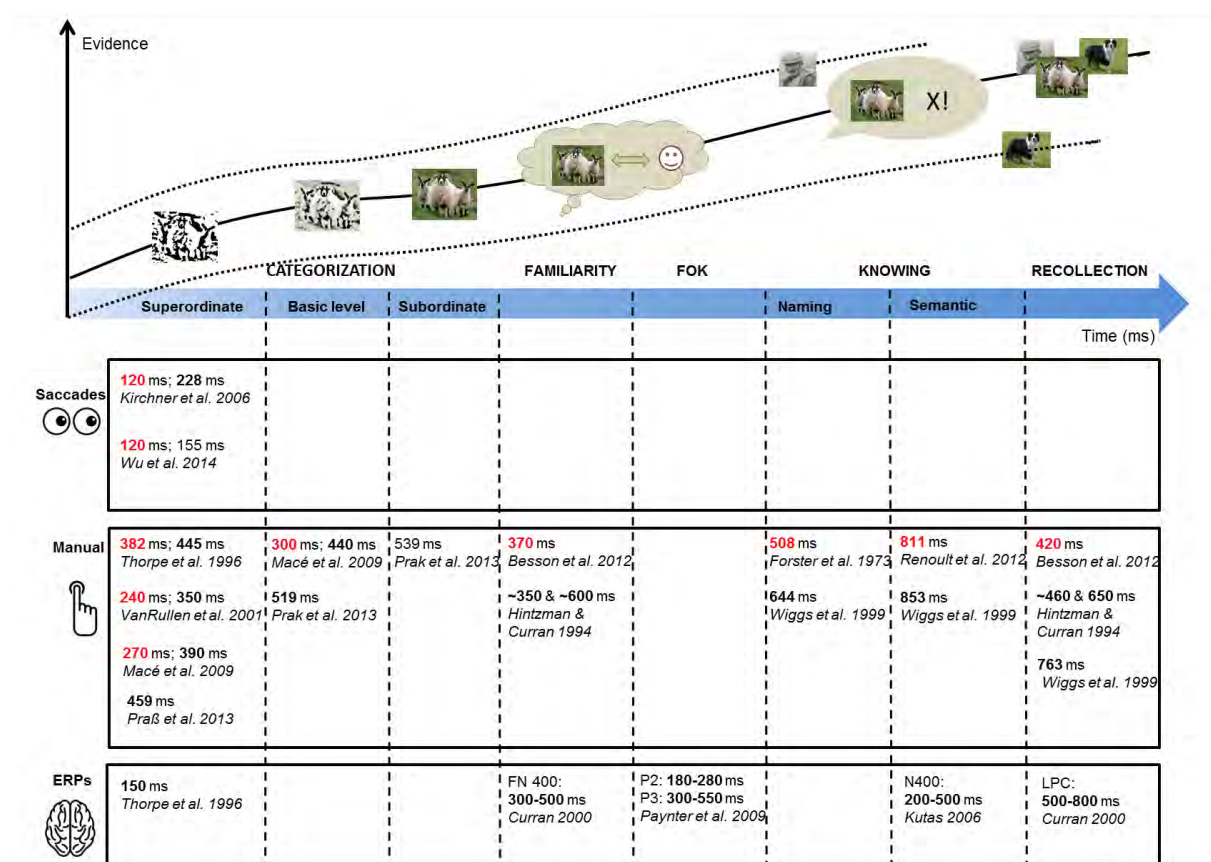


Figure 1.3 Decisions based on the evidence accumulation model.

Over time more evidence is accumulated and the categorization of the stimulus becomes more refined. The fastest timing (in red) and the average timing (in bold black) are shown for saccadic, manual and Evoked Response Potentials (ERPs) with the corresponding authors of the experiment (in italics).

3.1 Timing of early processes

3.1.1 Priming effects

Perceptual decisions do not necessarily rely on subjective awareness and can be simply made on subliminal perceptions. In a masking paradigm, a stimulus is flashed and quickly followed by another masking stimulus that makes the first stimulus invisible. The first stimulus is not subjectively perceived, yet it has some influence on further decision processes. This paradigm has been used in several similar experiments involving among others number comparison (Del Cul et al., 2007) or animal detection (Bacon-Macé et al., 2005). For example, DelCul et al. showed that even when participants reported that they had not seen a digit, they were still above chance at comparing that digit to 5. EEG was also recorded during the task and a direct comparison was possible between conditions in which participants reported having seen the stimulus or not. Until 250 ms, brain activity was broadly similar between the two conditions with activity originating from both occipital and temporal regions. In agreement with Bacon-Macé et al. (Bacon-Macé et al., 2005) brain activation increased monotonically during this early stage.

Interestingly, from about 270 ms they noted substantial difference between subjectively ‘seen’ and ‘unseen’ stimuli with strong activation in fronto-parieto-temporal regions for seen stimuli (Bacon-Macé et al., 2005; Del Cul et al., 2007) associated with a strong non-linear increase (sigmoid shape) for stimuli consciously perceived. This all-or-none measurement is correlated with subjective perception and means that perceptual awareness has been reached.

3.1.2 Categorization

When implicitly or explicitly perceived, a stimulus can be categorized at different conceptual levels involving some degrees of precision: superordinate level (e.g. animal), basic level (e.g. bird) or subordinate level (e.g. cardinal). While many studies agree that it takes more time to process a stimulus at a subordinate level than at the two others, there is a debate about whether the superordinate or basic level is processed first. Early work (Mervis and Rosch, 1981; Murphy and Smith, 1982; Rosch et al., 1976) was in favour of the basic level as the first categorization processed in the brain. However recent studies (Joubert et al., 2007; Loschky and Larson, 2010; Macé et al., 2009; Poncet and Fabre-Thorpe, 2014; Praß et al., 2013) have shown that in fact stimuli were first categorized at the superordinate level and ~ 50 ms later at the basic level. Unlike earlier works, the main difference was that categorization did not

imply any semantic processing and was mainly based on the analysis of visual input. These results collected on manual responses were confirmed in a very fast forced-choice saccade task (Wu et al., 2010). This study showed that 120 ms was enough for participants to detect the presence of an animal (superordinate) but not to determine if it was a dog (basic level).

Priming effects and categorization are the earliest processes that can be made without necessarily relying on subjective awareness. The results above are compatible with the suggestion that those decisions are made based on the evidence accumulation model that we described in the previous sections. Information accumulates over time and enables making refined decisions (Cichy et al., 2014). Interestingly, these early decisions correspond to the first stages of memory encoding. It has been shown that neurons in IT have already enough selectivity to make categorizations at the basic level (Cichy et al., 2014; Kriegeskorte et al., 2008). Subordinate categorization, which needs more expertise is supposed to rely on PFC (Mack and Palmeri, 2011). Interestingly, this hierarchical processing observed for visual inputs is preserved for the other sensory modalities such as audition (De Lucia et al., 2010; De Meo et al., 2015; Murray et al., 2006).

3.2 Timing of higher level processes

Whereas a lot of data exist in terms of speed of processing for early decisions far fewer studies have measured the timing of later processes. Nevertheless, several paradigms have been proposed to measure the differences between high level processes such as familiarity, Feeling of Knowing (FOK), semantic and episodic memory.

3.2.1 Familiarity versus recollection

The first method to quantify the speed of familiarity and recollection is the Speed Accuracy Tradeoff (SAT). In this paradigm, a word is presented and participants have to judge if it was already seen or not under different time constraints. By using the SAT, several studies (Boldini et al., 2004) have shown that familiarity is processed before recollection. However the SAT does not give the precise timing of the two processes. To do so, Besson et al., (Besson et al., 2012) proposed a new paradigm: the Speed Accuracy Boosting procedure (SAB) that has been used with visual stimuli. This method is similar to a go/no go task (press a button if the stimulus was previously seen/hold the response) but adds a time constraint: responses have to be made within 500 ms. Under such conditions, the fastest response occurred at ~ 370 ms. Using the same paradigm, this result was replicated for famous faces (Barragan-Jason et al., 2013). Interestingly Besson et al. 2012 also collected remember and know judgements

when participants were performing the SAB. They found that from ~ 370 to ~ 420 ms responses were only based on familiarity. Only later are responses possibly based on "remember" judgements. Those behavioral results are in agreement with EEG studies that separate familiarity and recollection into two respective components: 1/ the FN400 that occurs in the frontal region between 300 and 500 ms after stimulus onset and 2/ the Late Positive Component localized in the centro-parietal region between 500 and 800 ms (Curran, 2000).

3.2.2 Feeling of Knowing

The FOK is the feeling that we can retrieve an item without having it yet in mind. The "game show" paradigm (Reder, 1987) has been designed to measure participant's ability in judging information to be possibly retrievable. In a modified version (Reder and Ritter, 1992) participants were given 850 ms to decide for a strategy in an arithmetic problem. They had to decide if they will be able to solve it by retrieving from memory or by making a calculation. Under this strong time constraint, participants were not able to give the answer and could only rely on their 'gut feeling', but participants were good at evaluating which strategy was the optimal. Thus, FOK latencies are shorter than those for giving the correct answer. Several other studies have been performed (Koriat, 1995) with the same conclusion: FOK appears before the subjects can report the response. Unfortunately in FOK judgement tasks, time constraints are usually lenient and do not really offer the opportunity to measure when that process starts. Nevertheless it has been possible to track the time-course of FOK by EEG recording. By using a new version of the game show paradigm, (Paynter et al., 2009), two main components of correct FOK were found: a frontal P2 component between 180 and 280 ms after onset trial and a fronto-central P3 component between 300 and 550 ms. This shows that as early as 180 ms there is enough evidence to start response initiation.

3.2.3 Semantic and episodic memory

The Remember/Know paradigm (R/K in short) was first proposed by Tulving (Tulving, 1985a) to dissociate episodic from semantic memory which corresponded to remember and know responses respectively. However, in common experimental designs 'Know' responses are not semantically related and instead refer to a familiarity judgement (Gardiner and Richardson-Klavehn, 2000; Williams et al., 2013). Most of the studies using R/K/N paradigms are thus comparing familiarity and recollection processes. However, in an imaging study using Positron-Emission Tomography (PET), (Wiggs et al., 1998) directly compared semantic and episodic retrieval. In their study they found that naming of an object was

processed earlier than semantic and episodic information about the objects. However, surprisingly, reaction times were faster for retrieving information based on episodic details compared for semantic-based retrieval. This finding does not fit with ERP studies (Curran, 2000) that showed recollection is a late process that occurs between 600 and 800 ms (late positive component) and where semantic information is processed between 200 and 500 ms (N400). To our knowledge, no task has been designed to test and compare directly the speed of semantic and episodic retrieval in conditions where participants have to make very fast responses as in the SAB paradigm. We note that reaction times collected by Wiggs et al. are long compared to other studies (Forster and Chambers, 1973; Renoult et al., 2012). In this case, the contradictory behavioral finding from Wiggs et al. with ERP studies could be explained by the fact that their design was not encouraging participants to respond quickly enough (as is done in go no-go designs such as with the SAB) and that, as a consequence, the participants' decisions may have been 'contaminated' by other processes. Our evidence accumulation model makes a different proposal. As explained in Section 2.2, recollection corresponds to a sufficient amount of evidence accumulated for several items related to a hypothesis: dashed lines Figure 1.3 Depending on the strength of these distributions, retrieval of secondary items- that are related to the self or not- can occur earlier or later than the main one which is sought for.

Overall these studies depict memory encoding as a continuum of processes: categorization, familiarity, semantic and episodic information. The crucial point here is that at every point of this continuum, the information is encoded at the level it was processed. Retrieval of the information will be then computed on the evidence coming from the stored-representation of each of these levels. Therefore, if encoding was stopped at an early stage, retrieval will be only based on low-level properties of the stimulus. Whereas a small amount of evidence may be enough to make judgements based on a mere feeling of knowing, more information has to be computed to reach higher levels of retrieval.

4. Clinical application

By extracting the maximum a posteriori information after some time of search and computing the confidence on the selected option, our model seems optimal in generating reliable memory judgements. We have shown that critical regions might support these decisions and especially the rhinal cortices, the hippocampus and the (pre)frontal cortex.

Brain damage or lesions in these structures or simply momentary dysfunctions of their neuronal network might perturb the system and elicit inappropriate memory judgements ultimately leading to dementia. In the following we will demonstrate that our model can account for such disorders by focusing on Alzheimer's disease.

Clinical studies of Alzheimer's disease point to neuronal damage that first occurs in the MTL -with the entorhinal cortex as a primary site (Braak and Del Tredici, 2012; Gomez-Isla et al., 1996; Khan et al., 2014)- which then spreads out to neocortical regions and corresponds to an increase of dementia severity from mild amnesic impairment (MCI) to Alzheimer Disease (AD) (Niikura et al., 2006; Serrano-Pozo et al., 2011).

Numerous studies have shown that recognition memory is particularly impaired in AD patients but this is less clear for MCI individuals. The traditional view is that this reflects differences in the memory system affected by damage. It is also consistent with our accumulation of evidence proposal, even when considering the progression from less severe (MCI) to more severe (AD) dementia and the differences between two different types of recognition tasks that differ in terms of environmental support. As an example, people with MCI are not impaired on forced-choice recognition tasks but showed deficits on Yes/No recognition tasks, whereas AD patients are both impaired on Yes/No and forced-choice recognition tasks (Clark et al., 2012).

This discrepancy of performance in the MCI population parallels blindsight patients in perceptual domain. As a result of strong damage in the primary visual cortex, people with blindsight are unable to report seeing a stimulus when it is presented in their 'blind' hemifield. Nevertheless, they can still perform better than chance level if they are asked to discriminate two stimuli. It has been recently suggested that the overall neuronal activation during the presence of a stimulus might decrease after such lesions requiring the update of the Present/Absent criterion (Ko and Lau, 2012). This inability to shift the criterion in blindsight patients -which is the same as considering the absent class with a high prior value (J-R. King and Dehaene, 2014)- might explain patients' performance.

We propose that a similar explanation might account for MCI and AD patients to explain the results obtained in recognition tasks. In a Yes/No task the distinction between "Old" and "New" items is based on a "familiarity" criterion (see blue line, top diagrams, Figure 1.4) in a single dimension (one stimulus). Therefore an item would be considered as "Old" if its

posterior probability – which depends on both its prior probability and its likelihood – is larger than the posterior probability of encountering a new stimulus which here is equivalent to its prior. The posterior probability of an “Old” and “New” item are represented respectively by orange and grey circles. Mild cognitive impairment arguably arises due to the reduction of neural signals originating in the neuronal assemblies in the entorhinal cortex (see the middle top plot, Figure 1.4) that might be a crucial region in the placement of the “familiarity” criterion (see (Didic et al., 2011)). Because of this signal decrease, a shift in the familiarity criterion would be necessary to adjust the placement of this criterion to a less conservative value. This might result in the prior value decrease of new items, that is to say, a decrease in expecting new stimuli to appear. The position of this updated criterion is represented by the dashed blue line (middle top plot, Figure 1.4). As such, people with MCI will not be able to make this shift of prior and decisions might rely on a non-updated criterion that is now too conservative (middle top plot, Figure 1.4). As a result, people with MCI would have a trend to consider information that they already experienced as new events. This would support our hypothesis of considering that the MTL structures are directly involved in updating the prior value. However, because the hippocampus and neocortical neurons are still preserved, this overall decrease does not impact the ability of people with MCI to perform above chance level on a forced-choice task. In a forced-choice task the distinction between two items: here an “Old” and a “New” item is based on the “knowing” criterion (see orange line, bottom plots, Figure 1.4) in a two dimensional representation (two stimuli). Unlike in a Yes/No task the prior probability of a new item does not depend on individuals’ past history. This prior probability is the same for the “Old” and “New” items and is fixed at 0.5 in every trial of a forced-choice task. A decision in a forced-choice task would therefore rely on the comparison of the likelihood of the two hypotheses which requires less evidence than in a Yes/No task. In the case of more severe loss of neurons, (for example in AD patients), the loss of neurons in both the hippocampus and the neocortex could result in a large overlap of the two neuronal populations which would prevent their differentiation. As such, more severely amnesic people are unable to report having seen a stimulus before (top right plot, Figure 1.4), and their performance in a forced-choice is also impaired (bottom right diagram, Figure 1.4). Ultimately, if there is no capacity to represent two differing neural populations which map onto separate concepts, then memory performance in a forced-choice task will be at chance levels. Again, this explanation of memory deficits can be viewed in a dynamic way. The amount of evidence accumulated by more severely impaired patients (e.g. later stages of AD)

would stop at a very early stage and even before the first stage as depicted in Figure 1.2 whereas less impaired people (e.g. MCI) would be able to reach this first level.

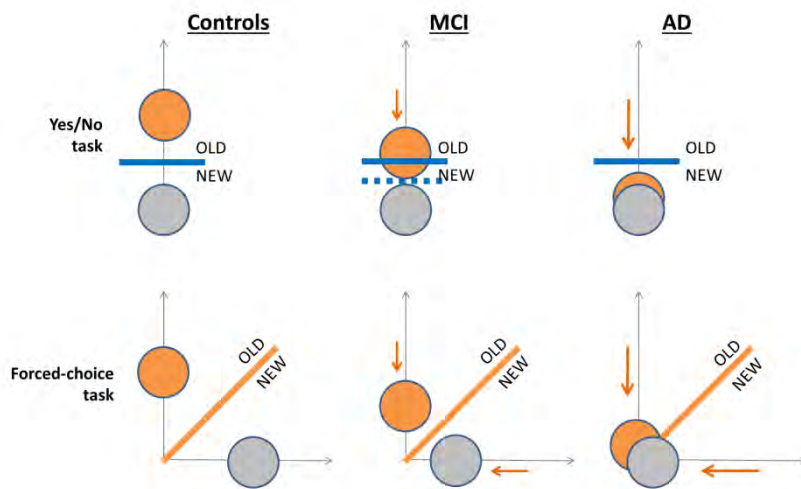


Figure 1.4 Two-dimensional representation of the distribution of the neuronal responses on a Yes/No judgement and a on a forced-choice task.

The neuronal distributions involved in a memory task are represented for controls, MCI individuals and AD patients. The decrease of neuronal signal is depicted by the orange arrows. Blue line represents the optimal familiarity criterion for normal neuronal activations whereas the dashed blue line corresponds to the optimal familiarity criterion in case of brain damage in MCI individuals. Orange line corresponds to the “knowing” criterion.

5. Conclusion

Perception and memory are not two distinct entities but need to be considered along a continuum. A dynamic Bayesian perspective is commonly used in the field of perception but until now has been poorly developed in terms of memory retrieval. This chapter is a first attempt at trying to define retrieval mechanisms according to that scheme. This model already fits experimental data and sheds some light on pertinent memory debates such as the dual versus single process debate and the consideration of confidence judgement in familiarity and remember responses. But more strikingly, by identifying two crucial parameters, the familiarity and knowing criterion, this model gives strong predictions about memory access that can be applied to clinical cases. One of the next challenges would be to find a way of manipulating these parameters and potentially bring back to mind memory traces that were thought to be lost.

By our view, the dynamic Bayesian approach developed in this review based on the interaction between a prior and a likelihood value is particularly relevant when considering the retrieval of recent memories that are associated with a high prior. However, recent or frequently rehearsed memories are not the only types of memories that remain in the brain. Memories of events/stimuli might stay dormant for years (Dudai, 2004), yet retrieval would still be possible. The probability of encountering these stimuli would drop to a value so close to zero that the prior value might not have any influence in the retrieval of such memories. We suggest that the retrieval of these long-lasting memories might only be supported by the similarity between the current input and the stored-pattern: the likelihood in the Bayesian formulation.

CHAPTER 2. From recent and active memories to remote and inactive ones

Abstract

We will argue that the notion of “activity level” might be more relevant to study memory retrieval and especially when considering (very) long-term memories. In addition to the time variable, the activity level of a memory takes into account the potential re-activations of a memory trace. It makes it easier the differentiation between old memories that are often re-activated and remote memories that were not re-expressed for a long time and that remain in a dormant state. To account for the retrieval of memories with various levels of activity we will use the same dynamic Bayesian framework that we largely developed in chapter 1. By considering that neurons might compute Bayesian inferences which results in the generation of spikes, we will show a straightforward link between memory activity and neuronal activity. This perspective would enable us to speculate about the maintenance of inactive memories over time.

Presentation chapter

In chapter 1 I have presented memory under the scope of a Bayesian framework. Memory retrieval would be the result of prior expectation together with the likelihood of a current input.

In a previous version of this chapter review I wanted to integrate a section on very long-term memories. Indeed, it is the main topic of my thesis. But as I already said, Simon was not very favorable to explain long-term retrieval using the Bayesian theorem. According to him, how could we retrieve very long-term memories that were associated with a very small prior? For memories that are inactive for years, this value must be almost zero. In the classical Bayesian formulation, the posterior probability is the result of the multiplication between the prior and the likelihood. If the prior is almost zero, there is almost no chance for retrieval. This was problematic because the goal of my thesis was to show evidence for very long-term memories.

I agreed with Simon: the prior value would be close to zero for inactive memories. But in this case, retrieval would only be based on the second factor, namely the likelihood. It took me months to be able to express it simply. During this time, I thought that it would be helpful to show that the retrieval of very long-term memories would be similar to Tversky's similarity asymmetry calculations as already shown by Tenenbaum and Griffiths (Tenenbaum and Griffiths, 2002). But it was even worse. I was so stuck that I decided to send an e-mail to Tom Griffiths. He replied right away and suggested me to read John Anderson's articles. Such a good advice! From these articles I eventually realized that this problem could be fixed very easily: the Bayesian formulation should be written using a logarithm scale where: $\text{posterior} = \text{prior} * \text{likelihood-ratio}$ becomes: $\log(\text{posterior}) = \log(\text{prior}) + \log(\text{likelihood-ratio})$. I should have been more careful about the use of the logarithmic scale when I read it first from Gold and Shadlen (Gold and Shadlen, 2007). This transformation changes a product into a sum and it is now easy to demonstrate that memories with a low prior value could be retrieved if the $\log(\text{likelihood ratio})$ reaches a certain threshold.

In addition, John Anderson's formulation of the Bayesian theorem relied on the notion of memory activity/inactivity. To me this was crucial because the main goal of my thesis was to demonstrate that we could retrieve memories that were left inactive for years. After months of thought, I finally gathered enough information to provide an explanation to how very long-term memories that were left dormant for years could still be retrieved. I explain it in this

second chapter. I also provide biological considerations about how Bayesian inferences might be computed in the brain, something that was lacking in chapter 1.

1. Introduction

The commonly held view is to categorize memories according to the duration between the formation of a memory and its retrieval. The short-term/long-term dichotomy is a classic example that can include several subtypes such as “very short-term memories”, “recent memories”, “very long-term memories”, “remote memories”. This perspective however shows limitations considering that memories can be revisited during the intervening period. In that sense, old-memories that are reactivated might share similar properties to new memories. To that extent, Lewis (Lewis, 1979) suggested that memories should be considered in terms of activity levels. Active memories might refer to newly created memories or to old memories recently revisited that can be easily disrupted by amnesic agents such as electroconvulsive shocks. Inactive memories might correspond to old memories that were not reactivated recently and that are not sensitive to amnesic agents. Accordingly, most of our long-lasting memories might remain inactive and immune to changes by staying in a “dormant” state (Dudai, 2004).

To be retrieved, memories need to be first available and then accessible (Tulving and Pearlstone, 1966). Memory availability relies on the persistence of the memory trace. If synaptic connections or neurons are lost, such availability might be disrupted and could lead to forgetting. The stabilization of these long-lasting traces occurs slowly through consolidation that might depend on the modality of presentation of a stimulus and its emotional content. However, availability *per se* is not sufficient to elicit retrieval. Indeed, retrieval is completed when the memory is re-accessed again at a specific time. The temporary inability to access information might be considered as a second explanation for forgetting. According to a recent study (Habib and Nyberg, 2008) the difference between accessible and temporarily inaccessible memories might be simply due to the level of neuronal activation. If too low, access might be impaired.

According to this view, all the memories that were consolidated and that were not reactivated for a long time might still be available. By staying “dormant”, inactive memories might therefore remain available. They might rely on neocortical structures that provide massive storage capacities with their 20 billion neurons. But as illustrated by their level of resistance to amnesic agents, one of the main differences between active and inactive memories might be the ease to access them again. Inactive memories might be hard to wake up from their dormant state compared to active memories that are associated with a high expectation value (prior value).

In this chapter we will show how consolidation allows the gradual formation of stable memories and their long-term availability. To account for the retrieval of memories with different activity levels we will use the dynamic Bayesian model that was developed in chapter 1 in the context of re-activations. In particular, we will demonstrate that for inactive memories, the prior probability might be so small that retrieval might be only elicited by specific cues from the environment. Although the Bayesian theory is usually perceived in an abstract way, we will suggest that it could be considered at the neuronal level and grounded in the overall brain architecture. This would enable us to suggest that in/active neurons might be the biological substrate of in/active memories.

2. Memory consolidation

Memory consolidation corresponds to the formation of a stable memory trace performed in two stages (Frankland and Bontempi, 2005). The first one called synaptic consolidation is fast (completed within a few hours after learning) and corresponds to the stabilization of synaptic weights in localized neuronal networks. The second one, much slower and called system-level consolidation refers to a change of brain structures supporting the memory trace.

2.1 Synaptic consolidation

2.1.1 Hebbian rule

“Neurons that fire together, wire together” (Schatz, 1992). This hebbian rule is well known in the neuroscience community and is a simplistic view of Hebb’s postulate (Hebb, 1949): “When an axon of cell A is near enough to excite a cell B and repeatedly or persistently takes part in firing it, some growth process or metabolic change takes place in one or both cells such that A’s efficiency, as one of the cells firing B, is increased” (p 62). According to this view, synapse strengthening would then require the cell A to fire before the cell B. The importance of firing temporal order in the change of synaptic weights was further detailed with Spike Timing Dependent Plasticity (STDP) models.

2.1.2 Spike-Timing Dependent Plasticity (STDP)

In the standard STDP models, only the synaptic inputs that fire in a time window close to a post-synaptic spike are modified: either strengthened or weakened. In a sense, STDP models can be considered as an extension of the Hebbian rule by adding the notion of synaptic

depression. Synaptic inputs firing before the post-synaptic spike are potentiated whereas synaptic inputs firing after are depressed (Bi and Poo, 1998).

2.2 Repetition: the key for neuronal selectivity and long-term storage

If a subjective representation contains strong emotional content, one exposure can be enough to lead to permanent storage (Buchanan, 2007; Reisberg and Hertel, 2005). In cases where the emotional content is lighter, repetition might trigger long-term memory storage in the neocortex, as already proposed by David Marr in the 1970s (Marr, 1970). Along this line, we suggest that repetition of the same stimulus will indicate per se that information might be useful and needs to be consolidated. This can be explained in terms of neuronal mechanism in which Spike Time Dependent Plasticity (STDP) rule is applied.

2.2.1 Repetition-suppression and repetition enhancement

Repetition might play a crucial role by means of 2 phenomena: repetition-suppression (for a review see: (Grill-Spector et al., 2006)) and repetition-enhancement (for a review see: (Tartaglia et al., 2015)). Repetition-suppression corresponds to a diminution of neuronal activity that can be easily measured at various scales: from single neurons to hundreds of thousands of neurons in fMRI voxels. This phenomenon is very robust and has been largely replicated through various experimental conditions. More subtle is repetition-enhancement that usually comes with repetition-suppression and that might only concern a very small amount of neurons.

Several theories have been proposed to account for repetition-suppression. The fatigue model suggests that the repetition of the same stimulus might lead to a general “fatigue” of the neurons involved with a decrease of response amplitude in the signal. The facilitation model explains this decrease by a switch of signal latency inside the neuronal population. These two theories however do not explain the phenomenon of repetition-enhancement.

The sharpening model accounts for both repetition-suppression and repetition-enhancement by suggesting that over repetition, the neuronal population involved in the encoding of the stimuli might get smaller and therefore more specialized. While the specific neurons might continue to fire the same way or even more strongly than when the stimulus was presented for the first time (repetition-enhancement), most of the remaining neurons that are not specifically coding for the stimulus might not fire again generating a general activity decrease (repetition-suppression) (see Figure 2.1). The sharpening model was proposed by Desimone (Desimone, 1996) in light of intracranial recordings in monkeys and then considered as a potential

mechanism for explaining priming effects (Wiggs and Martin, 1998). Importantly, repetition-suppression is not a passive phenomenon but might be directly caused by the inhibition from the small number of neurons that are becoming selective (Thorpe, 1990).

Sharpening is a good model for explaining the increased selectivity for stimuli that become highly familiar (Anderson et al., 2008; Freedman et al., 2006). However, it fails to explain the suppression in the most selective neurons after the repeated presentation of items that are mildly familiar (De Baene and Vogels, 2010; Li et al., 1993).

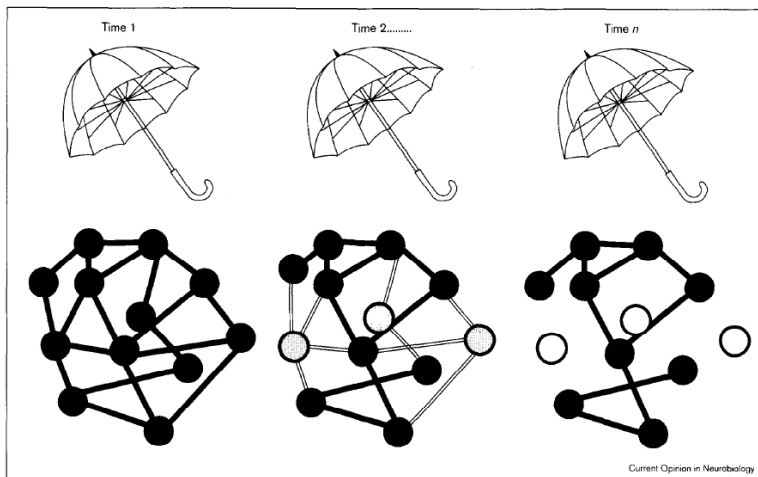


Figure 2.1 Illustration of the Sharpening model.

A large neuronal network (bottom of the picture) might encode a stimulus such as an umbrella (top of the figure) when it is presented for the first time. Over repetition, neurons that are not specific to the stimulus might not fire anymore and only selective neurons might be able to fire again. Therefore, the memory trace becomes more specific and smaller, that is, sharper (from (Wiggs and Martin, 1998)).

2.2.2 STDP and computational models

According to the STDP model, neurons that, by chance, fired when a stimulus was presented for the first time might strengthen their connections over the repetitions. Neurons that were not responding during the presentation of the stimulus or responding too weakly would be then inhibited.

Some recent computational studies have used a simplified version of the STDP rule, in which after a post-synaptic spike, all synaptic inputs are depressed, except the ones that have fired just before. As a result, neurons equipped with such simplified STDP-learning rule become selective to repeated patterns and only after a ten of repetitions (Bichler et al., 2011;

Masquelier and Thorpe, 2007). It is simply over repetition and through this competitive-STDP-based-learning that the neuronal network encoding a specific stimulus might become sharper (Masquelier et al., 2009) and might rely on cortical areas (Norman and O'Reilly, 2003).

2.2.3 Neuronal selectivity

Computational data reveal the need of a few tens of repetitions (defined as the repeat of patterns that were previously learnt (Wilson and McNaughton, 1994)) to achieve the formation of selective neurons.

Evidence for rapid selectivity was also shown in participants who incidentally learnt auditory noise patterns (Andrillon et al., 2015). The data in the Supplementary Information of the article clearly demonstrate that five presentations are enough to create a fully distinct ERP component. If such a rule applies for hippocampal and neocortical neurons, it would mean that neurons would naturally become selective to input patterns that occur repeatedly.

Synaptic weights in the hippocampus can be modified quickly after a few repetitions of the same stimulus but are short-lasting. To be stored in long-term, memories might be supported by synaptic changes in the neocortex. Long-lasting selectivity which might be specific to neocortical neurons would then explain the extreme long retention of stored-representation in these regions. Indeed, if selectivity is high enough, neurons would remain silent and in an inactive state until the stimulus they encode is presented again.

By reactivating a recent memory trace during 'off-line' stages, the hippocampus might speed-up the formation of new neocortical assemblies of neurons supporting the memory (Davis and Gaskell, 2009). Neocortical hyper-selectivity might also occur without its support but in such case it might require more repetitions from the environment as suggested by hippocampal lesions in patients (Freed and Corkin, 1988; Gabrieli et al., 1988; Verfaellie et al., 2000).

2.2.4 'Off-line' reactivations during consolidation

Reactivations of a new memory trace can occur spontaneously at rest, right after a new learning task (Buzsáki et al., 1983), or during sleep (Peyrache et al., 2009) and correspond to sharp wave-ripple (SW-R) complexes. These complexes are constituted by two components: large depolarizations called sharp waves and ripples which are brief bursts of fast oscillations. A disruption of these SW-R complexes directly impairs memory consolidation (Girardeau et al., 2009; Jadhav et al., 2012). During this phase, SW-R complexes in the hippocampus are

associated with slow wave oscillations in the prefrontal cortex allowing a dialogue between the two structures and the strengthening of the memory in the neocortex.

2.3 System-level consolidation

Storing of items and concepts in the hippocampus as they are subjectively perceived occurs alongside activations in neocortical regions and in the prefrontal cortex for example (Preston and Eichenbaum, 2013). However, because the hippocampus has limited storage capacities, consolidation models suggest that long-term memory traces might become independent of the hippocampus and will rely only on neocortical structures. Two main theories have been proposed to explain such phenomenon: the standard model and the Multiple Trace Theory (MTT).

2.3.1 Standard model

According to the Standard model (Squire and Alvarez, 1995) both semantic and episodic memories might become independent of the hippocampus after consolidation. This was proposed first in light of patients with retrograde amnesia who were impaired in retrieving information that was encoded before hippocampal damage. While memories encoded just before the damage were lost, the older were the memories and the more they were preserved. This temporal gradient, extending for several years, was supposed to reflect the increasingly independence of declarative memories from the hippocampus.

2.3.2 Multiple Trace Theory

The Multiple Trace Theory was developed a few years later, as an alternative to the radical Standard model (Nadel and Moscovitch, 1997). If indeed, semantic long-term memories might not need the support of the hippocampus, the retrieval of episodic memories might still rely on hippocampal functions.

Both theories share the assumption that this structural reorganization comes along with memory reactivations. According to the standard model, reactivations would only affect cortical regions whereas for the multiple trace theory, every reactivation might generate a new memory trace in the hippocampus.

3. Post-consolidation reactivations, prior and accessibility

The availability of a memory trace can be defined as the preservation of its biological trace (Tulving and Pearlstone, 1966). It is an intrinsic property of a neuronal pattern that has been

stabilized through consolidation. As such we consider that any consolidated memory might last for a lifetime as long as its biological trace remains intact. Therefore post-consolidation reactivations might be not necessary for a memory to be available. Importantly, post-consolidation reactivations might not necessarily result in explicitly retrieving the memory trace. If encoding was restricted to low level activations without involving the hippocampus, reactivations might remain implicit. Indeed, studies using meaningless auditory noise have shown that it is possible to retrieve information that does not need any reactivation during the intervening period and that is not necessarily explicit (Agus et al., 2010; Viswanathan et al., 2016). But what we usually observe is that over time, it gets harder to retrieve past information: the information is still present but its accessibility, the ability to “find” the available memory (Tulving and Pearlstone, 1966) becomes restricted. Then the role of post-consolidation reactivations would be to facilitate the access of a memory trace and as it means literally, to turn inactive memories into active and accessible memories.

3.1 Memory activity in a Bayesian perspective

The idea to formulate memory activity in a Bayesian framework was already proposed by John Anderson in the late 90’s through the idea of rational analysis (Anderson and Matessa, 2003; Anderson, 1995). In his rational analysis of memory Anderson postulated that the “activation-level” (A) of a memory (i) corresponds to the summation of a “base-level” (B) and a “contextual-priming”.

Mathematically, “base-level” is the log prior odds and reflects to what extent a memory is likely to be used again;

“contextual-priming” is the log likelihood ratio that takes into account cues (j) from the environment;

“activation” level is the log(posterior odds) which is an inference about which memory to make available given the context.

The activation level of a memory A_i can be therefore written as follows:

$$A_i = B_i + \sum_j W_j S_{ji}$$

with B_i the “base level” activation of the memory, W_j the weighting of the cues and S_{ji} the strength of the association between the cue j and the memory.

The prior probability depends on the time since a memory has been activated and on the number of times it was encountered. If a stimulus was presented once, the probability to encounter it again might be high right after its presentation, but would rapidly decrease over

time unless it is presented again. The repeated presentations of a stimulus and their spacing in time might be critical to update the prior value. In the next section we explain how reactivations of a consolidated memory might help their retrieval.

3.2 Updating the prior value

The prior probability distribution of a specific memory reports to what extent we expect to encounter again the memory. When a memory was recently formed its probability of re-occurrence is high, and therefore its prior. As already discussed in chapter 1, in addition to being encoded in the neocortex, these memories with high prior values might be supported by the hippocampus. After some time, the said recent memories become old and are associated with low expectations and might preferentially rely on neocortical structures. We suggest that the role of post-consolidation reactivations might be to slow down this shift of prior value: every time information is retrieved from memory its prior might be updated to a strong value and as such being supported again by hippocampal structures.

Interestingly, in a dynamic Bayesian framework, only the likelihood value varies over the decision time. Decision is the result of the accumulation of the likelihood ratios over time and the prior is considered as the likelihood ratio at a time $t = 0$. As such, the same different levels of retrieval -from implicit retrieval to full recollection- can emerge for memories that have not been re-experienced for years (inactive memories) or for memories recently encountered (active memories).

We will use this framework to explain how retrieval might emerge considering memories with different levels of activity.

3.2.1 Frequently reactivated memories

When memories are very often reactivated they acquire a strong prior that make them more distinguishable from other memories and therefore more easily retrievable (Figure 2.2.A). This is why it's normally easy to answer the question: 'Who's the current president of the Unites States?' This memory is fresh enough to be able to report the name without the need of additional support. Active memories can correspond to recently learned information but also to (very) long-term memories frequently rehearsed.

3.2.2 Memories with scarce reactivations

When memories are reactivated from time to time, the prior value drops-off and it is more difficult to get access to them (Figure 2.2.B). To compensate the low prior value, some

external support can be given to restrict the dimensional memory space and facilitate retrieval. For example, try to remember *who was the last former UK prime minister?* For most people a cue will be at least necessary to get the name back to mind.

3.2.3 Remote memories that were not reactivated

A few exposures to a given stimulus can be enough to allow the formation of a stable and long-lasting memory trace in the neocortex (Bichler et al., 2012). In many cases however these robust memories might not be rehearsed again and are referred in the literature as inactive (Lewis, 1979) or dormant memories (Dudai, 2004). These memories can be implicit (Mitchell, 2006) but might also involve higher levels of retrieval such as the ability to recall a name or to judge an item as familiar.

For memories that have not been reactivated since then, the expectancy to encounter that memory and therefore the value of the prior is very low (Figure 2.2.C). If we want to turn them active again, a strong environmental support is necessary. Therefore the more the overall new experience will match the old one, in terms of stimulus and overall context associated to it (Reed, 1931; Smith et al., 1978) the more will be the chance to bring dormant memories back to mind. This could be done by showing the same stimulus again among a limited number of options such as in a forced-choice task and/or by using stimuli rich in content such as audiovisual clips.

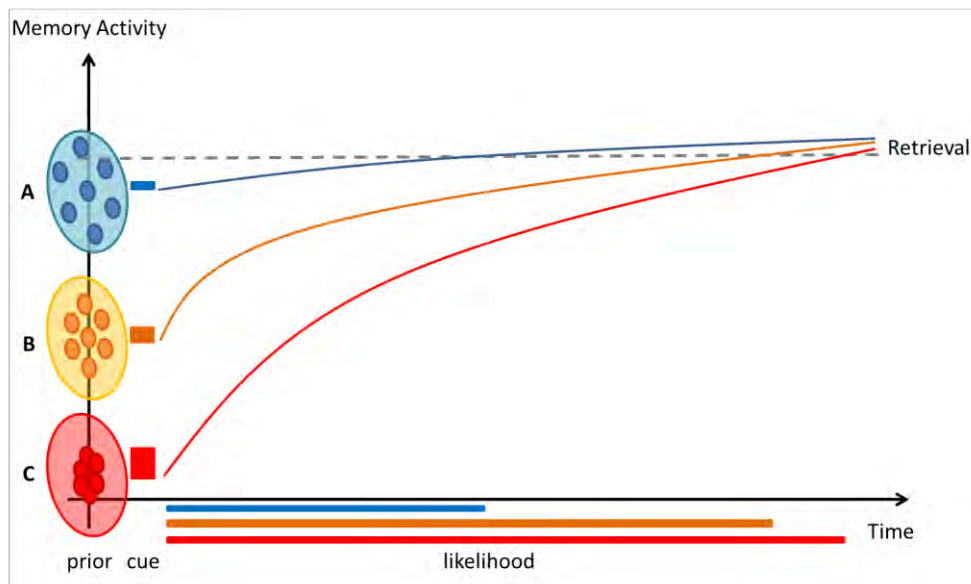


Figure 2.2 Retrieval from inactive to active memories.

Often-reactivated memories get strong prior values that make them well differentiated from each other (A). Weak cues (blue square) or no cue at all are needed to elicit the retrieval of those active memories. Barely-reactivated memories get lower prior values that make them

less distinguishable from the other memories (B) and therefore need stronger cues to be retrieved (orange square). Memories that have never been reactivated after consolidation get very low prior values (C) and need strong cues to elicit the retrieval of those inactive memories (red square). The amount of activity for eliciting retrieval in a given task is represented by the dashed line.

Post-consolidation reactivations are useful to turn inactive memories into active and accessible ones but do not seem to play a major role in the availability of the memory per se. Moreover, it has been shown that the reactivation of a memory can sometimes be associated with its reconsolidation (Dudai, 2012). Reconsolidation is in fact considered as an update of previous learnt information (Alberini, 2005; Dudai and Eisenberg, 2004; McKenzie and Eichenbaum, 2011) and can lead to its distortion (Sara, 2000). As a consequence, if we want to keep a memory intact for decades, reactivations do not seem necessary and might be rather harmful.

4. Silent neurons for dormant memories: experimental evidence

So far, little is known on the neuronal substrate of dormant memories. Indeed, most of the studies that worked on very long-term memory focused on the impact of subsequent reactivations of a stable memory trace. Reactivations strengthen the memory trace and spread it out to a larger distributed neuronal circuit involving both the hippocampus and the neocortex (Moscovitch et al., 2006). We suggest that the formation of a stable memory trace is associated with the creation of a small and specific population of neurons in the neocortex that will stay silent until the reactivation of the memory.

4.1 From STDP to dynamic Bayesian inferences

4.1.1 Neuronal level

It has been recently proposed that Bayesian inference might not remain abstract and could be simply explained in light of the STDP and neuronal mechanisms (Nessler et al., 2013). According to this view, the prior corresponds to the intrinsic excitability level of a neuron which directly depends on its firing activity (Daoudal and Debanne, 2003). Therefore, active memories that correspond to memories with a high prior value might be supported by highly excitable neurons. This means that by having ion channels opened in their axon, dendrites and soma, such neurons would easily generate an action potential (the output of the neuron) in presence of specific stimuli. Interestingly, the logarithm of the likelihood ratio, might

correspond to the sum of the inhibitory and excitatory post-synaptic potentials (I/EPSP) in the neuron. The values of these post-synaptic potentials directly depend on the synaptic weights. In that sense, one I/EPSP might correspond to one piece of evidence for making the neuron fire or not. The interaction between neuronal excitability (prior value) and post-synaptic potentials (likelihood ratio over time), might elicit the firing of a neuron. As a response, each spike generated by a neuron might correspond to the calculation of the posterior probability at a specific time and the whole spike train, showing the accumulation of evidence over time.

4.1.2 Hierarchical structure

Neurons in the brain are part of a network which occupies a specific position in the overall architecture. Through this hierarchical structure, the prior probability of neurons in low-level structures might directly depend on the likelihood ratio computed in the upper-level structures (Friston, 2005). This view fits the idea that priors are constantly learnt through experience and shows that most of the feedforward models currently used to model recognition might not be sufficient to explain memory processes (Serre, 2016).

4.1.3 Temporal order coding

By firing neurons convey information to the system. Usually neurons do not generate one spike but rather a train of potentials. Information might be therefore encoded in the overall firing rate of the neurons according to a “rate-based” coding view or in the temporality of the spikes according to “spike-based” models (Brette, 2015) where for example the respective order of each spike might be informative (Thorpe, 1990). Indeed, it has been shown that the firing rate code hypothesis might present some limitations when considering very fast decisions and that a rank order consideration might be more suitable (Gautrais and Thorpe, 1998; Van Rullen and Thorpe, 2001). Such very fast processing might rely on the first spike(s) and would not require waiting until the end of the whole train of potentials as for a rate-based coding. Interestingly, this rank order theory might fit well our dynamic Bayesian evidence model at the scale of the neuron and could explain the spiking variability found in individual neurons during intracranial recordings. The accumulation of spikes might correspond to the accumulation of evidence in time that could stop very early if neurons are in high level of excitability (high prior). Therefore, one spike generated by a neuron might be sufficient to convey enough information. Conversely, if no spike is generated or if the spikes occur very late, the chance to get the information to be processed might be very small. As an example, the spiking variability recorded in the Medial Temporal lobe of an epileptic patient

over eight trials, for four presentation conditions and for two stimuli (Quiroga et al., 2008b) is presented in Figure 2.3. Interestingly, when the patient was not able to recognize a picture (red trials), the neurons selective to the picture did not fire at all or generated a first spike very late after the stimulus presentation. Overall, note the “long” latency of the first spike for the two neurons which indicates their selectivity to the stimuli (Mormann et al., 2008).

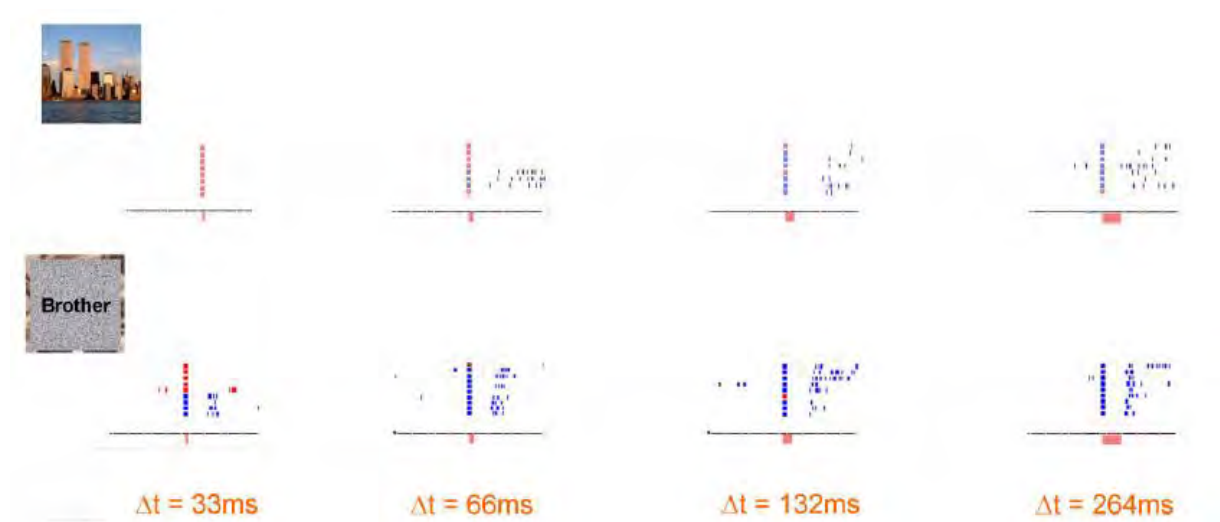


Figure 2.3 Activations of neurons during the presentation of a specific stimulus.

Neuron specific to the World Trade Center (top of the figure) or to the brother of the patient (bottom of the figure, covered for privacy) recorded in the entorhinal cortex and in the hippocampus respectively. The eight diagrams correspond to the raster plots of the neurons where each line represents the firing of the neuron (1 bar = 1 spike) for each trial (1 line = 1 trial). Each stimulus was presented eight times (eight lines for each raster plot) and for four various durations that are shown by the light red square at the bottom of each raster plot (from 33 to 264 ms). Trials where the pictures were recognized by the patient are in blue and trials where the pictures were not recognized are in red. (Adapted from (Quiroga et al., 2008b)).

Unlike many other spiked-based models, the rank-order theory relies on the asynchronous firing of the neurons: most of the information might be conveyed by neurons that fire first in a winner-take-all-fashion (Thorpe, 1990).

4.1.4 Oscillations

At a more global scale, the asynchronous firing of neurons might generate a decrease of power in low-frequency brain oscillations such as in the alpha-beta range and an increase in higher frequencies such as in the gamma range (Hanslmayr et al., 2012). Decrease in alpha-beta power has been shown to be important for both the encoding (Jiang et al., 2015) and the retrieval of memories (Düzel et al., 2003). However, and as pointed out recently (Hanslmayr and Staudigl, 2014), many studies have tried to correlate specific brain oscillations to memory processes resulting in a varieties of conclusions and even sometimes the emergence of contradictory results. Therefore, a better view would be to consider brain oscillations in terms of level of processing, - a view that is perfectly coherent with our dynamic Bayesian model- rather than using generic terms such as memory encoding and retrieval.

4.2 Inactive neurons for inactive memories

4.2.1 Local versus distributed representation

How is a mental representation encoded in the brain? Although simple, this question is still a matter of debate and no consensus has been reached. The main view considers that mental representations are supported by a large distributed network that would involve millions of neurons. On the other side, others have argued that a more restricted population of neurons would underlie our percepts and memories (Barlow, 1972; Bowers, 2002; Quiroga, 2012; Thorpe, 2011, 1989).

When one thinks of this second alternative, it is in direct association with the concept of “grandmother cell”. This term was proposed by Jerry Lettvin in 1969 (unpublished notes, see his letter written to Barlow in 1991 (Barlow, 1994) and an excerpt from this letter (Gross, 2002)) and refers to the idea that one cell could encode a concept such as a “grandmother”. This concept was not seriously considered by the whole community if not completely mocked but probably because of misinterpretation (Bowers, 2009). Indeed, and as pointed by Bowers, a localist coding does not imply that rich mental representation might be only encoded by redundant single neurons. While it might be true for specific concepts such as a grandmother, the integration of several concepts into a specific neuronal network might be the key for encoding our rich visual representations and memories. Instead of grandmother cells, the use of concept cells would be therefore more appropriate (Quiroga, 2012; Quiroga et al., 2013; Reddy and Thorpe, 2014). As for words used in language, these concepts cells would be able

link low-level features to create a simple but distinct and meaningful representation. The combination of these concept cells in a specific order might be able to encode and retrieve specific episodes, individual sentences of our life. This localist view might be referred as sparse distributed rather than fully localist or fully distributed.

4.2.2 Sparse coding in the medial temporal lobe

According to the sparse distributed point of view, encoding and retrieval of even rich contextual information might be supported by the activation of a limited amount of neurons. Sparse coding is indeed found in the medial temporal lobe and in the hippocampus in particular (Quiroga et al., 2008a, 2005; Rolls, 1989; Wixted et al., 2014) where only a maximum of 1% of the cells would be active when a stimulus is presented (Waydo et al., 2006). This means that a neuron would fire for about a couple of hundreds of distinct concepts. Sparseness in primates' or in humans' hippocampus is not new and shows extreme analogies with place cells found in rodents' hippocampus (O'Keefe and Dostrovsky, 1971).

4.2.3 Dark matter in the neocortex?

While active memories might be supported by highly excitable neurons, inactive memories might rely on silent neurons. Sparse coding is not restricted to the Medial Temporal Lobe and can be found in the neocortex (Graham and Field, 2006) including the prefrontal cortex (Abeles et al., 1990). A recent finding (Barron et al., 2016), has shown that it is through an excitatory-inhibitory synaptic balance that cortical memories might stay silent and that its disruption is able to re-express these dormant memories.

For many years, studies focused on the detectable differences of brain activity elicited by the firing of neuronal populations but interest is growing in the consideration of a kind of neocortical "dark matter" (Shoham et al., 2006) corresponding to the whole set of cortical neurons with very low firing rate at a given time. As with dark matter in the universe, it is possible that most of the human neocortex could be made up of silent neurons, consistent with theoretical calculations that consider that neocortical neurons might fire 0.16 spikes a second on average (Lennie, 2003). Although experimental evidence in humans has demonstrated the existence of silent and highly specific neurons in the medial temporal lobe (Ison et al., 2011) which are directly implicated in the formation and retrieval of recent memories (Gelbard-Sagiv et al., 2008; Quiroga, 2012), silent neocortical neurons might be the neuronal substrate of remote memories (Thorpe, 2011).

5. Conclusion

In this chapter we extended the dynamic Bayesian framework that was developed in chapter 1 to the retrieval of inactive memories. Using the notion of level of activity we were able to close the gap between this theoretical framework and biological mechanisms. In particular, the firing of neurons would be the result of Bayesian inferences. As such, inactive memories would be encoded by inactive neurons with a low-level of excitability. Yet possible, the retrieval of these dormant memories might be therefore difficult and would only rely on the stimulation of these neurons. We suppose that the best way to elicit the re-activation of these inactive neurons would be to reconstitute with the most details the scenarios in which the stimuli were presented. In such a way, some concept cells at the basis of the memories could fire again and potentially in a coherent manner to elicit the retrieval of a whole dormant episode. Experimental investigations are now required to confirm this hypothesis.



- **PART 2 – EMPIRICAL FRAMEWORK -**



Presentation of the behavioral experiments

In the M4 project, I was the “archeologist” of the team. Like many others, during my childhood I was passionate about remote civilizations. I was fascinated by the fact that scientists could tell stories from tiny remains that they found in specific spots. From a piece of fabric, from a text fragment, they could reconstruct a particular moment of some individuals’ life. Fifteen years later and in a different way, I had to play this role for which no name exists. My field of investigation was not a lost civilization, it was remote memories. In this neuroscientific version of the archeologist, I did not go to an exotic country to find remains of remote memories. Instead, I stayed in the south of France, in my lab in Toulouse and also in Marseille which was my farthest venture (presented in chapter 4).

Of course, I was not the first one and some renowned neuroscientists have already investigated this field. All together these studies have shown that memories can be maintained for years and potentially for a lifetime. However, two main factors remain unclear: do memories survive the passage of time because they are reactivated from time to time? In other words, can memory persist for decades in a latent way? The second question refers to the minimal exposure duration. What is the smallest number of exposures and the shortest time required to create a memory trace that could be maintained for decades?

To answer these questions, the critical point was to test remote memories that individuals could not have reactivated since the initial period of exposure. The easiest way would have been to find participants that would agree to see stimuli in the lab and to come back let’s say 20 years later to test their memory. However my thesis had to stick within a 3-year limit... To overcome this issue we had to test individuals’ memory for stimuli they had experienced already and control that the memory of these stimuli had not been reactivated since then. Such stimuli would have been presented in a previous experiment run by colleagues such as the colored-drawings used in chapter 4 or acquired from personal experience like TV series that individuals watched during their spare time and which is the topic of this next chapter.

CHAPTER 3. Waking up buried memories of old TV programs

Abstract

Although it has been demonstrated that visual and auditory stimuli can be recalled decades after the initial exposure, previous studies have generally not ruled out the possibility that the material may have been seen or heard during the intervening period. Evidence shows that reactivations of a long-term memory trace play a role in its update and maintenance. In the case of remote or very long-term memories, it is most likely that these reactivations are triggered by the actual re-exposure to the stimulus. In this study we decided to explore whether it is possible to recall stimuli that could not have been re-experienced in the intervening period. We tested the ability of French participants ($N = 34$, 31 female) to recall 50 TV programs broadcast on average for the last time 44 years ago (from the 60's and early 70's). Potential recall was elicited by the presentation of short audiovisual excerpts of these TV programs. The absence of potential re-exposure to the material was strictly controlled by selecting TV programs that have never been rebroadcast and were not available in the public domain.

Our results show that 6 TV programs were particularly well identified on average across the 34 participants with a median percentage of 71.7% ($SD = 13.6$, range: 48.5% - 87.9%). We also obtained 50 single case reports with associated information about the viewing of 23 TV programs including the 6 previous ones. More strikingly, for 2 cases, retrieval of the title was made spontaneously without the need of a four-proposition choice. These results suggest that re-exposures to the stimuli are not necessary to maintain a memory for a lifetime. These new findings raise fundamental questions about the underlying mechanisms used by the brain to store these very old sensory memories.

Presentation chapter

Without a doubt, testing participants' memories of old TV programs was not the easiest to set up. First of all, I had to find TV series from the 60's and 70's which had not been re-broadcast and that were not available on websites or video-tapes. In short, I had to find French audiovisual archives that were lost from the public domain for about 40 years. Why this period? Because at the end of the 70's, people started to use videocassette recorders and it would have been impossible to control for the absence of re-exposure after the original broadcast. In addition, it was a great opportunity to test memories that had lasted a lifetime. I had the chance that the French Institut National de l'Audiovisuel from Toulouse gave me the opportunity to access to its archives. From their database I was able to get short excerpts of 50 TV programs that would constitute my test videos.

My material gathered, I was ready to test the memory of individuals who had a television in the 60's and 70's in France. My colleague Sophie and I visited all the senior clubs of the city, the playground of this first experiment. In our hands, we had only a laptop which held the experimental set up that Simon, Nadège and I designed together.

From the beginning, I knew that in such experimental conditions it would be difficult to get a lot of results, if we were to get any results at all! A main limit was I could not know which TV series were watched by a particular individual. As a consequence, I could only report when a participant was able to retrieve some information about a TV program but I could not evaluate the truth of the memories or judge which memories were forgotten. In practical terms this means no percentage of retrieval. Despite these important limitations it was a formidable opportunity to test real-life memories of stimuli that were acquired in ecological conditions. These stimuli were not just words or pictures but were short animated clips combining auditory and visual components. I knew that if results we got, it would be only thanks to the richness of these stimuli.

1. Introduction

As adults, we all have memories of sounds and images that were formed decades ago. For people who are now in their 70s and 80s, these memories are part of their very long-term memory also called remote memories. Memories can be semantic if they reflect general knowledge such as the ability to retrieve a movie title or episodic if they involve the recollection of a unique specific event in space and time: for example remembering a particular day when someone watched a movie on their first date (Tulving, 1985a). When related to the self, both semantic and episodic memories create autobiographical memories that are specific to each individual. Therefore, autobiographical memories can involve generic facts about personal past events ("I used to watch my favorite TV series with my brother") that are not always episodic (Levine et al., 2002; Piolino et al., 2002; Tulving et al., 1988).

Important neuronal reorganizations are required to create long-lasting memories and involve two consolidation stages (Frankland and Bontempi, 2005). The first one which is called synaptic consolidation refers to the stabilization of synaptic weights of a new memory in localized networks. This process is fast and can be completed within a few hours after learning. The second one, involving system-level consolidation, is much slower and corresponds to a change of brain regions that support the memory. Both the hippocampus and the neocortical structures are initially involved in supporting declarative memories (semantic and episodic memories). However, within several months after learning, theories suggest that declarative memories might become independent of the hippocampus. This would concern semantic and episodic memories in light of the so-called standard model (Squire and Alvarez, 1995) or simply semantic information regarding the Multiple Trace Theory (Nadel and Moscovitch, 1997). Note that the debate is still not closed between these two theories.

Given the practical difficulties involved, only a small number of studies have tried to test recall decades after the acquisition phase. Such studies have looked at memories concerning old classmates (Baird et al., 1975), information learnt in school (Baird, 1984; Conway et al., 1991) or even TV programs (Squire and Fox, 1980; Squire and Slater, 1975; Squire, 1989). Although the stimuli used are different, the results follow the same trend: recall drops quickly over the first 6 years and then levels off for several decades.

As mentioned by the authors, in such studies, one parameter that is hard to control fully is potential reactivations of the information during the intervening period which might explain the extremely long retention of these memories. It has been shown that memory reactivations can be triggered spontaneously during periods when the processing of sensory input is very low ('off-line states') such as during sleep (Diekelmann and Born, 2010; Peigneux et al.,

2004; Wilson and McNaughton, 1994) or wakefulness (Foster and Wilson, 2006; Karlsson and Frank, 2009; Oudiette et al., 2013; Peigneux et al., 2006) as well as during ‘on-line states’ when subjects are actively retrieving memories (Gelbard-Sagiv et al., 2008; Nyberg et al., 2000; Tayler et al., 2013). While reactivations during ‘off-line states’ are critical for the consolidation of a new memory (Gais et al., 2000; Girardeau et al., 2009) and are very frequent in the first few hours after learning (Eschenko et al., 2008; Ribeiro et al., 2004), their probability of occurrence is likely to decrease exponentially over time (Frankland and Bontempi, 2005; McClelland et al., 1995), leaving the memory in a dormant state (Dudai, 2004; Lewis, 1979; Sara, 2000). This suggests that the reactivation of these dormant or inactive memories might occur spontaneously, especially during ‘on-line’ states when sensory input is strong enough to elicit memory retrieval, that is, during the re-exposure to the stimulus or when it is mentally evoked. For about 15 years now and since the discovery of Nader et al. (Nader et al., 2000) it has been shown that reactivations of stable and consolidated memories through specific stimulus re-exposure might trigger re-consolidation processes. During this temporary unstable stage, memory traces are updated (Dudai, 2006) and strengthened (Lee, 2008; Moscovitch et al., 2006) and as a result, become more accessible and are less vulnerable to decay (Gisquet-Verrier and Riccio, 2012). Accordingly, it might be natural to think that remote memories could be maintained via subsequent re-consolidations that are triggered by specific stimuli, even if these re-consolidations are scattered in time. This leaves open the question of whether it is possible to retrieve very long-term memories for stimuli that have not been re-experienced for decades and that are left in a dormant state in the absence of any subsequent reactivations.

In this study we tested the ability of French people to recall a selected set of 50 TV programs that were originally broadcast with between 6 and 120 episodes between the late 50’s and early 70’s. These TV programs have never been rebroadcast and are not available in the public domain so we can be certain that the stimuli have not been seen or heard since the original broadcast. Furthermore, the participants reported that they had not thought about them for years, making it unlikely that they would have involuntarily reactivated the memories (Rasmussen and Berntsen, 2009).

By sorting participants performance over the confidence in their response, we found 6 TV programs with a percentage of correct identification (median: 71.7%, SD =13.6, range: 48.5% - 87.9%) that was significantly higher than for younger participants. Interestingly, these 6 TV programs were part of a set of 23 videos for which we collected single case reports with associated information about the viewing at the time of broadcast. Whereas identification was

mainly performed by selecting the correct title from four propositions, in two cases, the presentation of short excerpts of old opening themes was able to trigger spontaneous naming of the title. Overall our data suggests that visual and auditory memories can indeed be retrieved even when they have been buried for decades.

2. Material and Methods

2.1 Participants

The protocol was approved by the INSERM Ethical Evaluation Committee (CEEI) and all participants gave their written informed consent before experimentation.

Thirty four subjects, all French (31 female; range = 52-92y, median age = 77y, mean age = 79y, SD = 7.6) with normal or corrected-to-normal vision and audition participated in the study. All of these participants were recruited and tested individually in senior citizen clubs. Before starting the experiment participants were invited to respond to a questionnaire to give personal details about their TV habits. They were asked to say roughly when their household first had a television and how many hours a day they typically watched TV at the time of the test (Supplementary information: table S1). The participants reported first having a television sometime between 1956 and 1970 with half of them having access to a television before 1964 (SD = 4.1) and said that they currently were watching TV for an average of 3.5h a day (range 0.5-9h, SD = 2.1). After the experiment, participants gave their feedback about the TV programs they were able to recall. None of them reported thinking about the ones which had never been rebroadcast. The overall cognitive abilities of 25 out of the 34 older participants were also assessed using the Mini-Mental State Examination (MMSE) (Folstein et al., 1975) based on the French GRECO consensual version. No main deficit was found for any of the participants (mean score 27.6 out of 30, SD = 1.9, range: 24-30).

Thirty-four younger participants, ($t(66) = -32.0$, $p < 0.001$), all French (23 female; range = 21-40y, median age = 26y, mean age = 27y, SD = 5.5) with normal or corrected-to-normal vision and audition participated in the same study. Younger participants reported having access to a TV set from between 1976 and 1999 (mean year = 1992, SD = 5.9) and watching TV for 1.2h a day (range = 0-5h, SD = 1.5).

2.2 Stimuli

Audiovisual clips (size: 640*480) were presented on a gray background at the center of a laptop screen placed in front of the participant (Hewlett Packard EliteBook, screen resolution: 1366*768). The clips used a 1s count-down followed by a 7s opening theme of a TV program.

The clips generally displayed the first seven seconds of the opening themes without any text. In total 72 audiovisual clips were shown to each participant. Fifty of the clips were **test videos** that were composed of the opening theme excerpts of French TV programs which had never been rebroadcast and were not available in the public domain. With support from the French Audiovisual National Institute (INA), the videos were collected directly from an up-to-date database which has been collecting information on audiovisual programs broadcast since 1947. Videos were selected because they seemed original and because they did not change from the first to the last episode. The number of episodes was known for 36 of the 50 clips, and ranged from a minimum of 6 episodes to a maximum of 120, with a mean of 23.7. One of the programs was broadcast in the late 50's, thirty in the 60's and nineteen were first broadcast in the 70's. The average year of the last episode broadcast was 1970. The remaining 22 clips were **famous videos** of the opening themes of well-known TV programs which in some cases were still on air and were not necessarily French (Supplementary information: table S2). They were supposed to be easy to recognize and to keep participants' interest during the task.

2.3 Task

The experiment involved 72 trials with each video clip being presented only once. Each trial started with the presentation of an 8-second video clip (Figure 3.1). During the video or after its display, participants' recall responses were collected as follows: 1/ Does this TV program look familiar to you (Yes/No)? 2/ If yes, can you name the title of this TV program? (Free title naming). If not, please choose the related-title from the four propositions in the forced-choice (4-FC). The four propositions included the correct title, a lure (title of another TV program) and two foils (fabricated titles). The lures and foils were selected to be as plausible as possible. Propositions were displayed in alphabetical order. 3/ Rate your confidence in your response on a five-point scale (1: 'Not sure at all'; 2: 'A bit sure'; 3: 'Fairly sure'; 4: 'Very sure'; 5: 'Completely sure'). Five questions were also asked when participants reported being familiar with the TV program in order to get associated information: 1. What day(s) of the week did you watch this TV program? 2. Around what time of the day: in the morning, in the afternoon or in the evening? 3. How old were you at that time? 4. Did you like this TV program? and 5. Give as much information as you can about this TV program (e.g.: Who did you usually watch this TV program with? Have you watched a lot of episodes? What details could you give about the characters? etc.). Short breaks were made every ten trials. The experiment lasted about one hour and was programmed with Psychopy (Peirce, 2007).

In addition to the 34 participants tested, 34 younger participants performed the same task to make sure that the four titles proposed in the 4-FC would have the same probability of being chosen by naïve subjects who could only rely on semantic information to make their choice. Although it would have been possible to use age-matched controls, this was not the case here. We chose to test younger participants so that we could be certain that all the test videos were new to them. This would have been more difficult to control for older participants even if they said that they did not have a TV set in their house during that period, because at that time, it was relatively common to watch TV in someone else's home.

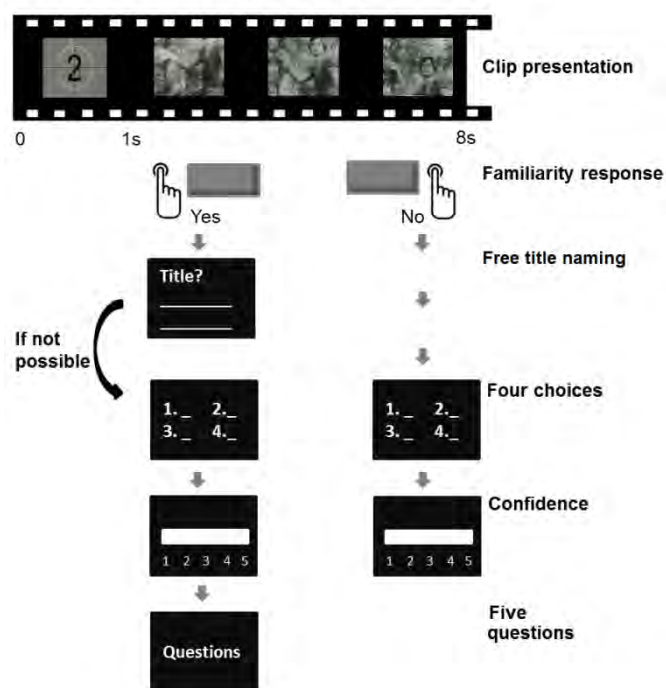


Figure 3.1 Experimental design.

Audiovisual clips of the TV programs were shown on a computer screen. Participants decided whether the program was familiar using a button press response. If familiar, participants attempted to give the title by free naming. Alternatively, they were given a four-proposition forced-choice (4-FC). The confidence level in the response was assessed on a five-point scale. For the familiar videos, five questions were asked to get associated information about the TV program.

3. Results

3.1 Participants' performance for the famous TV programs

During the presentation of the audiovisual clips or after their display, participants were first invited to decide whether the TV program was familiar or not. The older participants reported

that they were familiar with 56.5% (SD = 24.8) of the 22 famous audiovisual clips that were presented.

When participants said that a video was familiar, they were invited to name the title of the TV program directly (free title naming). On average, each famous video was spontaneously and correctly named by nine out of 34 older participants (27.5%, SD = 19.4, range = 0 - 25). This shows that the free naming of a title was high for the 22 famous TV programs and was always associated with a high confidence level (average = 4.5 up to 5, SD = 0.5). Overall, a median number of 6 (SD = 4.3) famous videos were correctly named spontaneously by the older participants, which was significantly lower than the 12 (SD = 4.9) correct titles reported spontaneously by the younger participants for the same famous videos (Wilcoxon signed-rank test: $Z = -4.18$, $p < 0.001$).

If participants found that it was too difficult to retrieve a title spontaneously, they were asked to select a title from four propositions (4-FC). When the older participants reported that they were familiar with the famous TV programs, the percentage of correct responses in the 4-FC was high (86.6%, SD = 21.7) and was not different from the younger participants (91.5%, SD = 14.9). The 4-FC was also directly proposed to the older participants who said that a video was not familiar. Overall, 68.8% (SD = 20.0) of the responses for famous clips were collected using the 4-FC including both familiar and unfamiliar responses. The older participants' performance was high when they had to choose the title of famous audiovisual clips (74.7%, SD = 15.5) and was not different from the younger participants (72.8%, SD = 12.9).

Interestingly, we found that the average percentage of identification for the famous videos was significantly correlated with the older participants' average confidence on the response ($r = 0.71$, $p < 0.001$, Pearson's linear correlation coefficient). The same effect was found in the younger participants ($r = 0.79$, $p < 0.001$, Pearson's linear correlation coefficient).

Overall, the older participants' performance was high for the 22 famous videos (median = 81.8%, SD = 9.7, range: 54.5 - 95.4%) showing that they could correctly perform the task. However, with a median of 90.9% (SD = 9.1, range: 54.5 - 95.4%) the younger participants' performance on the same videos was even better (Wilcoxon signed-rank test: $Z = -3.37$, $p < 0.001$). Such high levels of performance are explained by the fact that these famous TV programs are still very present in the media or even on TV. To what extent are the older participants able to identify old TV programs that have not been re-experienced for decades? We now address this question.

3.2 Older participants' performance for the test TV programs

3.2.1 *Single case reports for familiar test videos*

On average, the older participants reported that they were familiar with 15.2% (SD = 13.1) of the test videos.

- **Free title naming for test videos**

From the 50 test videos presented to 34 older participants, the titles of two TV programs were spontaneously recalled (free title naming). Although this percentage is very small (0.1%) it is still above the null score obtained by the 34 younger participants that were expected not to know the TV programs. Interestingly, these two free title namings were associated with a medium/high level of confidence in addition to associated information about the TV program. The following is a detailed description of the two responses:

On the first free recall, participant 5 (Supplementary information: table S1) said “Balzac”, which was close to the correct title “Un grand amour de Balzac”. Her confidence level was 3 out of 5. She was then asked to give details about her memories: “I was watching it on Saturday afternoon. I was sixty. Yes I’ve always loved sentimental and historical TV serials. I remember cousin Bette. I watched a lot of episodes.”

Actual facts: Episodes were broadcast in 1973 (she was 49 years-old, she is now 90) every Thursday at 9 pm. Seven episodes of 52 minutes were screened.

The other participant, participant 22 (Supplementary information: table S1) said “Camember” for “Les facéties du sapeur Camember”, with a confidence response at 4 out of 5.

The participant then reported: “I used to watch it on Sunday evening. I was thirty five. I liked it but I did not watch it too often. The whole family used to watch it.”

Actual facts: Episodes were broadcast in 1965 (he was 28 years-old, he is now 77), every day, except on Sunday. The TV program had 50 episodes lasting 5 minutes each. Initially the program started at 8.30 pm but switched to 9 pm from the 20th episode.

- **4-FC for familiar test videos**

When participants reported to be familiar with a TV program but could not name it spontaneously they were asked to select a title from four propositions (4-FC). Overall the percentage of responses was 23.7% (SD = 24.8) for the correct titles, 43.8% (SD = 27.2) for the lures and 14.6% (SD = 16.7) and 17.9% (SD = 16.1) for the 2 foils (fabricated titles). For familiar famous videos, 9.0% (SD = 13.9) of the responses were attributed to lures and 4.4% (SD = 17.9) and 0% (SD = 0) to the two foils. This revealed that the older participants were

significantly biased toward the lures for the test videos that were associated with a familiarity judgment, an effect that was not present for the famous videos (two-way ANOVA, ranked data: videos: $F(1, 236) = 23.06$, $p < 0.001$, titles: $F(3, 236) = 60.81$, $p < 0.001$, videos * titles: $F(3, 236) = 46.42$, $p < 0.001$, post-hoc comparison using the Tukey-Kramer test).

Although the older participants were not above chance level on average across the 50 test videos we decided to analyze participants' performance for each TV program. Indeed, given the design of our experiment, we do not know whether the participants watched the whole 50 test TV programs and to what extent. We found that for 23 out of the 24 test TV programs that were correctly identified on the 4-FC associated information about the TV program was provided. On average these videos were reported 2.2 times ($SD = 1.4$) by 23 participants with a maximum of six times for "Un grand amour de Balzac". Overall, the older participants were then able to report associated information for 2.8% ($SD = 3.3$) of the 50 old TV programs. Those responses were followed by a mean confidence level of 2.1 ($SD = 0.9$) which was significantly lower than the confidence for the correct free namings (mean: 3.5, $SD = 0.7$) but significantly higher than the confidence for the correct responses in the 4-FC for test videos judged unfamiliar (mean: 1.6, $SD = 0.5$) ($H(2) = 10.8$, $p < 0.01$, post-hoc comparisons using the Tukey-Kramer test). These 48 reports in addition to the 2 reports associated with free namings are given in the Table 1. Here are a few examples:

Report 8 (Table 1): After the presentation of the TV program "Poker d'as", participant 13 (Supplementary information: table S1) reported that she was familiar with the TV program and chose the correct title on the 4-FC with a confidence rate of 2 out of 5 ("a bit sure", average confidence for all test TV programs: 1.3, $SD = 0.4$). Then she reported that she liked the TV program and watched it during the week, in the evening, when she was 30-35. She added that she used to watch it with her husband but without the children and remembered having seen several episodes.

Actual facts: "Poker d'as" was a one-season TV series of 26 episodes broadcast in 1973 from Monday to Friday at 8 pm. The participant was 37 at that time and was 78 when tested.

Report 46 (Table 1): Participant 16 (Supplementary information: table S1) reported she was familiar with "Animal Parade" and correctly identified it in the 4-FC. Her confidence rate was 3 ("medium sure", average confidence for the 50 test TV programs: 1.4, $SD = 0.8$). Then she reported watching it during the week and at weekends in the afternoon. She could not remember how old she was at the time but she remembered that she liked it and watched it with her children.

Actual facts: “Animal Parade” was a single-season youth TV program broadcast from the 14th to the 25th of February (week and weekend) in 1972 at 7.30 pm.

Report 45 (Table 1) is interesting because of the 21 famous and test videos considered as familiar by participant 7 (Supplementary information: table S1), she reported that there were only two TV programs she did not like: “30 millions d’amis” an animal TV magazine and “Que ferait donc Faber?”, a comedy series broadcast in 1969. We found that there had been a large controversy following the broadcast of “Que ferait donc Faber?” with a lot of criticism from the French newspapers and the viewers. Other details reported by participant 7 (who was 91 when tested) and in particular the day and time of broadcast or her age matched the actual facts.

Among the 34 participants that were recruited we had the opportunity to test a couple individually on the same day: participants 25 and 30 (Supplementary information: table S1) who have had a TV set since 1964. Interestingly, participant 25 reported being familiar with “Poker d’as” and chose the correct title in the 4-FC with a confidence rate of two (“a bit sure”, average confidence for the 50 test TV programs: 1.0, SD = 0.1). He said that he used to watch it on weekdays during the evening when he was fifty. He added that he liked the TV program and watched it with his family but not too often. This was when they were still in Tarn-et-Garonne (*a department in the South of France*) a bit before 1976 so when he was 36 (report 7, table 1). Interestingly by recollecting this last piece of information, participant 7 corrected himself concerning his age at the time of broadcast. And indeed, this TV program was broadcast on weekdays at 8 pm in 1976 when he was 33. Unlike her husband participant 30 did not report any familiarity with “Poker d’as” but picked the correct title from the 4-FC with a confidence rate of 3 out of 5 (“medium sure”, average confidence for the 50 test TV programs: 1.5, SD = 0.7).

A detailed analysis of the older participants’ performance shows that the number of famous and test TV programs for which associated information was reported ranged between 1 and 23 (mean = 12.8, SD = 5.3) and were negatively and linearly correlated with participants age ($r = -0.49$, $p < 0.05$, Pearson’s correlation coefficient). The three male participants who were tested (participants 12, 22 and 25) were ranked in the top half.

Concerning the test TV programs, 10 of the older participants were able to give associated information for at least two TV programs, 12 others for only one TV program, whereas the remaining 12 participants were not able to give associated information for any of the test TV programs including the youngest and oldest participants (participants 19 and 24).

These limited examples suggest that it is indeed possible to recall memories of TV programs that were not re-experienced for decades. Moreover our data suggest that retrieval can be associated with subjective familiarity, some confidence in the response and with information about the viewing of the TV programs. Although it is hard to verify fully the accuracy of the associated information reported by the participants (such as the fact that participants liked the TV program or used to watch it with specific relatives), some information such as participants' age or the days and time of broadcast could give an idea of the accuracy of the memory. However, as suggested by the last report (report 50), the memory of old TV programs might not always be associated with direct familiarity judgment. In the following section we analyzed the data on the 4-FC by including the TV programs considered as not familiar.

	Report (n°)	Participant (n°)	Confidence (/5)	What day?	What time?	How old?	Age during broadcast	Did you like it?	Other details
Un grand amour de Balzac (1973) <i>Sentimental serial</i> 7 episodes Thursday, 9 pm	1	5	3	Saturday	Afternoon	60	49	Yes, I always loved sentimental and historical TV serials	I remember cousin Bette with Balzac. I watched a lot of episodes
	2	17	2	Every day	Afternoon	50	37	Yes	I watched it alone, a lot of episodes
	3	14	1	Weekdays	Evening	50	36	Yes	I watched it with my husband, rarely
	4	8	3	X	Evening	75	47	Yes	I watched it alone
	5	2	2	X	Evening	35	43	Yes	I watched it alone
	6	29	2	Weekdays	Evening	X	43	X	I have never watched it
Poker d'as (1973) <i>Spy serial</i> 26 episodes from Monday to Friday, 8 pm	7	25	2	Weekdays	Evening	50	33	Yes	I watched it with the whole family, not too often, before 1976 in Le Tarn Et Garonne (<i>French department</i>), I was 36
	8	13	2	Weekdays	Evening	30-35	37	Yes	I watched several episodes with my husband but without the children
	9	11	1	Monday	Evening	40	35	Not much	I watched it with my husband, very rarely, maybe one episode
	10	23	1	X	X	X	30	Not much	I watched it alone
L'Homme de l'ombre (1968) <i>Police drama</i> 30 episodes from Monday to Friday, 7.40 pm	11	12	2	Weekdays	Evening	43	26	A bit	In the bar
	12	23	1	X	Evening	37	25	A bit	Alone
	13	4	3	X	x	45	24	x	I have never watched it, only heard it
	14	15	2	X	x	45	29	Not much	X
Frédéric le gardien (1965) <i>Adventure</i>	15	17	2	Weekdays	Evening	26	29	X	I watched it a few times with my husband

series 24 episodes From Monday to Sunday, Time: <i>unknown</i>	16	8	2	X	Evening	75	39	Yes	X
	17	4	3	X	X	45	21	X	I have never watched it but I heard it in the bar
	18	20	2	X	X	X	34	X	X
Bayard (1964) <i>Youth series</i> 13 episodes Every Thursday, 6 pm	19	13	2	Wednesday or Thursday	Afternoon	40	28	Yes a lot	The children liked it a lot, we watched all the episodes
	20	4	2	X	We were in the bar	20	20	X	I remember hearing the music
	21	15	2	X	X	35	25	No	I watched it with my husband
Champion (1964-1966) <i>Game show</i> Episodes: <i>unknown</i> Tuesday, 8.50p m	22	17	2	Every day	Evening	50	28-30	Yes	I watched it with my husband, we watched a lot of episodes
	23	13	1	X	X	40	28	X	I maybe watched it with my husband
	24	15	2	X	X	X	25	Not much	X
Courte Echelle (1974) <i>Youth TV program</i> ~100 episodes From Monday to Saturday, 6.35pm	25	8	3	X	Afternoon	70	48	Yes	I watched it often and alone
	26	32	5	X	X	X	33	Yes	I don't know if the whole family watched it but I watched it every time
	27	24	3	X	X	X	52	X	X
L'arche de Samsong (1972) <i>Youth TV program (mini musical play)</i> 37 episodes From Monday to Saturday, 3.20pm	28	30	1	Weekdays	Evening	30	34	Not very much	I watched it a bit with the whole family
	29	14	1	X	X	X	35	X	I have never seen it but the music reminds me of something
	30	13	1	X	X	X	36	X	X
Le train bleu s'arrête 13 fois (1965-1966) <i>Police drama</i> 13 episodes Friday every 2 weeks 9.35pm (or 22.35 pm for the 4th episode)	31	14	1	Saturday	Afternoon	35	28	Yes	I watched it occasionally with the children, when we were in the house in Toulouse
	32	30	2	Weekdays	Evening	35	27	Yes	I watched it with the whole family, occasionally
Les facéties du sapeur Camember (1965) <i>Adventure series</i> 50 episodes Every day except Sunday 8.30pm for the 1st to the 20th episode	33	22	4	Sunday	Evening	35	28	Yes	The whole family used to watch it but not too often
	34	8	3	X	Evening	60	39	Yes	I watched it alone
Teuf-Teuf (1968-1969) <i>Youth game show</i> ~111 episodes	35	17	3	Wednesday	Evening	30	32	Yes	I watched a lot of episodes with the children
	36	4	2	X	X	I was young	24	X	I only listened to the music, I have never

From Monday to Saturday, about 6pm						or older			watched it
Vol 272 (1964) <i>Drama series</i> 13 episodes Every Sunday, 7.25pm	37	14	1	Weekdays	X	X	27	X	I have never watched it
	38	4	3	X	X	X	20	X	It was a long time ago, I heard it in the bar
Miss (1979) <i>Detective series</i> 6 episodes Day ? Time?	39	12	2	Weekdays	Evening	63	37	Yes	I occasionally watched it in the bar
	40	23	2	X	X	35	36	A bit	X
En direct avec (1966-1968) <i>Political TV program</i> Monthly Day? Evening	41	30	3	Weekdays	Evening	30	28-30	No	I watched it a few times, with the whole family
Encore un carreau de cassé (1960-1961) <i>Youth game show</i> Episodes, day and time: unknown	42	12	2	Wednesday	Evening	37	18	A bit	I was in the bar, I watched it a lot
La boîte à malice (1979) <i>Youth game show</i> Episodes? Day? Afternoon	43	11	2	Wednesday	Afternoon	40	41	No	I watched it with my children when we were in Montauban (<i>French town</i>)
Le roi qui vient du sud (1979) <i>historical serial</i> 6 episodes, Thursday, 8.30pm	44	17	2	Weekend	Afternoon	30	43	Yes	I watched a lot of episodes with the whole family
Que ferait donc Faber? (1969) <i>Comedy series</i> 8 episodes, Thursday, 9.40pm	45	7	2	Weekdays	Evening	50-62	46	No	We watched it by chance. We watched a few episodes when we were in Castres (<i>French town</i>)
Animal Parade (1972) <i>Youth series</i> 12 episodes Every day, 7.30pm	46	16	3	Weekdays and weekend	Afternoon	X	41	Yes	I watched it with the children
Candice ce n'est pas sérieux (1969) <i>Comedy series</i> 20 episodes Every day, 1.20	47	13	2	X	Afternoon or evening	40	33	I think so	I watched it with my children and my husband. I don't know if i watched a lot of episodes or not

pm									
Commandant X (1962) <i>Spy serial</i> 10 episodes, Day: unknown, 8.50pm	48	27	1	X	Afternoon	30	22	Yes	48
Alexandre Bis (1974) <i>Spy serial</i> 6 episodes, Thursday, 8.50pm	49	33	2	X	Evening	About 50	49	Yes	X
Objectif Demain (1979-1981) <i>Scientific TV program</i> Episodes and day: unknown, 8.45pm	50	4	2	X	Evening	45	35	X	I watched it without paying much attention
La vérité sur l'espionnage (1967) <i>Adventures series</i> 13 episodes Monday, 10.30pm	0	14	1	X	X	X	30	X	I have never watched it

Table 3.1 Collection of 50 reports for the test videos judged familiar by the older participants.

The 2 reports in orange (report 1 and 33) were made after spontaneous retrieval of the correct title. The other 48 reports were given after correctly choosing the title in the 4-FC. The columns: “What day?”, “What time?”, “How old”, “Did you like it?” and “Other details” correspond respectively to the responses to the 5 questions asked to get associated information about the TV programs. The column “Age during broadcast” was the actual age of the participant during the broadcast of the TV program.

3.2.2 Single case reports for familiar test videos

Most of the older participants responses were collected via the 4-FC: 97.4% (SD = 3.5) for the test TV programs. On average, the older participants were 24.4% (SD = 5.7) correct in identifying the title of a test video, which is similar to the younger participants: 22.9% (SD = 5.3) that were expected to perform at chance level in this task (25%). It is important to notice that in our study and as opposed to classical recall experiments we do not know whether all

the 50 test videos were watched by every participant and to what extent. We therefore analyzed the older participants' performance on the 4-FC for every audiovisual clip.

Interestingly, we found that the average percentage of identification of the test videos was significantly correlated with the average confidence on the response ($r = 0.37$, $p < 0.01$, Pearson's linear correlation coefficient) which was not the case for the younger participants ($r = 0.06$, Pearson's linear correlation coefficient). As shown in Figure 3.2, by sorting the average performance of each test video over the average confidence in the response, six titles were particularly well identified by the older participants when they used the 4-FC, namely: "Les Facéties du Sapeur Camember" (16 participants out of 33), "Teuf-Teuf" (20 participants out of 32), "Vol 272" (26 participants out of 33), "Frédéric le Gardian" (24 participants out of 34), "Un Grand Amour de Balzac" (29 participants out of 33) and "Courte Echelle" (24 participants out of 33). The median percentage of correct identification for these six TV programs was high: 71.7% (SD = 13.6, range: 48.5% - 87.9%) and significantly different from the 33.8% (SD = 20.6, range: 8.8% - 61.8%) obtained by the younger participants from the same six TV programs (Wilcoxon Signed-rank test: $Z = 21$, $p < 0.05$). Note that all of these titles were also found in the reports presented in Table 1, including "Les Facéties du Sapeur Camember" and "Un grand Amour de Balzac" which were spontaneously named.

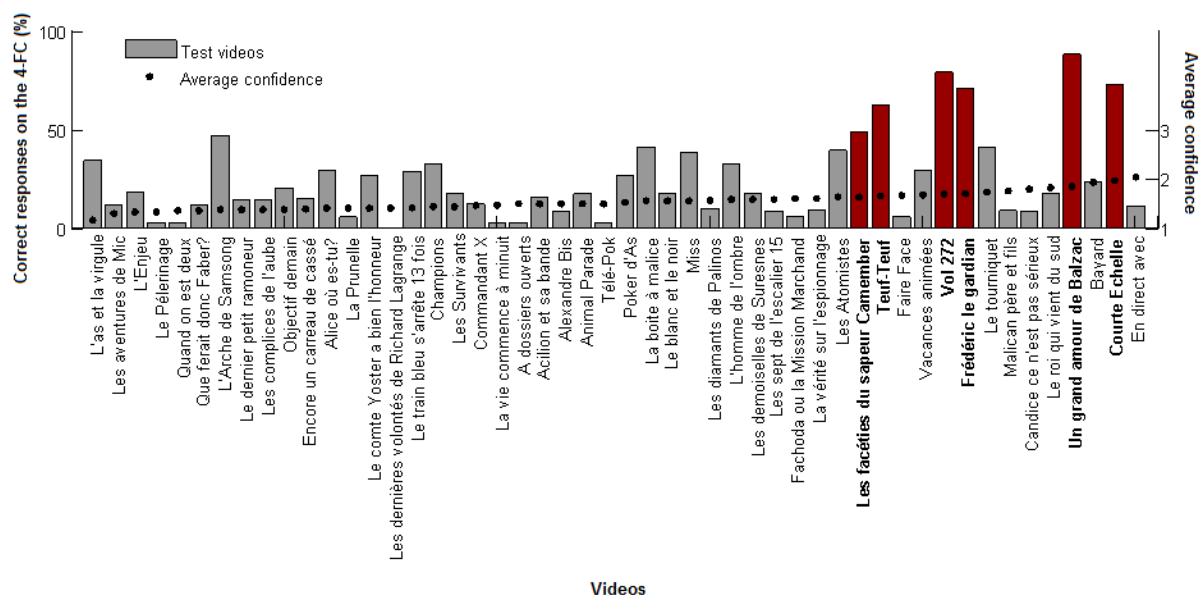


Figure 3.2 Performance in the 4-FC for the 50 test TV programs.

Videos were sorted according to the average confidence in the response given by the older participants in the 4-FC for videos judged familiar or not. The older participants' performance for the six TV programs are shown in red: "Les facéties du sapeur Camember",

“Teuf-Teuf”, “Vol 272”, “Frédéric le gardian”, “Un grand amour de Balzac” and “Courte Echelle”.

3.3 Influencing factors

In this experiment several factors might explain the variability in the older participants' performance for the test videos. This might be due to variations between the audiovisual clips or to individual differences.

In particular it would be interesting to find out if there was something about the 23 TV programs out of 50 that were associated with some contextual information, compared with the 27 that failed to work with anyone. This could be explained by the number of broadcasts, the overall duration of a TV program (episode duration * number) or the time interval since the last episode was broadcast. However, for each of these three factors we did not find any difference between the 23 TV programs recognized and the other 27.

Another source of variability might concern the audiovisual content presented during the 7s of each clip. In particular, some of the videos were more static than others. We therefore decided to count the number of different scenes defined as a specific action in space and time for each of the test videos (mean = 1.6, SD = 1.1, range: 1-7). We did not find any difference between the number of scenes displayed in the 23 test videos that were recognized compared to the other 27.

The older participants' performance for the 50 test videos might also be related to inter-individual differences. In particular we tested for the impact of five factors: age, cognitive abilities (given by the MMSE score), the year of TV set acquisition, the number of hours participants currently watched TV and the performance obtained for the 22 famous TV programs. We found that the percentage of correct responses for the 50 test videos was negatively and linearly correlated with participants' age ($r = -0.39$, $p < 0.05$, Pearson's correlation coefficient) and with the year of TV set acquisition ($r = -0.40$, $p < 0.05$, Pearson's correlation coefficient). However, we did not find any correlation between participants' performance for the 50 test videos and these following three factors: 1/the time the older participants spent watching TV when tested, 2/the older participants' performance for the famous videos, 3/ the older participants' overall cognitive abilities. For this latter factor, only 25 participants were tested.

4. Discussion

This study shows that recalling a complex audiovisual stimulus is possible decades after the original exposure, under conditions where reactivation in the intervening period is very unlikely. The absence of re-exposure to the stimuli was strictly controlled by testing TV programs that have never been rebroadcast and are not available in the public domain. However we cannot completely rule out the possibility that reactivation might occur during exposure to content that was indirectly related to these TV programs. Nor can we totally exclude involuntarily reactivations that might happen during sleep or wakefulness (Rasmussen and Berntsen, 2009; Rubin and Berntsen, 2009). Indeed, the stimuli used in this study are rich in terms of content and could elicit semantic associations with other items that are not directly related to the TV programs for example. Further experiments should use abstract/meaningless or simple stimuli to rule out this hypothesis. Nevertheless, given the absence of physical re-exposure to the stimuli and the fact that participants reported not having thought about these TV programs for years, the probability of occurrence of these reactivations might be very small when considering the retention interval between the time of broadcast and the recall. This gives enough support to consider that the memories of these old TV programs were dormant or in an inactive state.

Interestingly, all of the information retrieved concerned generic information about the test TV programs. They could be personal memories (participants' age, whether they liked the TV program or not, who they used to watch the TV program with and when) or not (title or characters' name of the TV program). The amount of personal details retrieved from autobiographical memory varied but even for the most detailed reports, no reference to a specific or unique event was mentioned. Several factors could explain participants' inability to retrieve any episodic autobiographical memory: the age (Rubin and Schulkind, 1997) and long retention interval between watching the TV programs and doing the test (Piolino et al., 2002) as well as the contextual similarity in which the episodes of the same TV program were watched. In the latter case, each episode of a TV program would be a new episode from which participants would create a generic pattern stored semantically (Neisser, 1981).

The retrieval of semantic autobiographical memory was associated with familiarity together with some confidence in the response. This shows that familiarity might be a property of the semantic system, a recent view that emerged from lesion studies (Vargha-Khadem et al., 1997). If this is the case, the recognition of a TV program was based on a judgement of previous occurrence performed consciously by the participant (Mandler, 1980). However, and as shown by the six TV programs that were identified with better than chance level

performance over the 34 participants, most of the recognized test TV titles occurred without any familiarity judgment and with low confidence levels. In this case, title recognition might only be due to an effect of perceptual fluency of which the participants were unaware of (Jacoby and Dallas, 1981; Verfaellie and Cermak, 1999; Voss et al., 2008). Participants might select a title because they could process it more readily given the video presented. Judgments based on perceptual fluency alone might not be the most optimal (Jacoby and Dallas, 1981) but could explain the bias towards lures for the test TV programs.

In our experiment, 23 out of the 50 test TV programs were able to elicit the retrieval of semantic autobiographical memories. We would have expected that a difference in the number of original broadcasts would have an impact on the retrieval of personal contextual details. However we did not find any difference between these 23 TV programs and the other 27. Indeed, as shown by reports 1 to 5 and 45 (Table 1) it seems that only a few exposures to a TV series or its opening theme (report 6, Table 1) can be enough to create a stable memory trace. We also thought that the time elapsed since the last broadcast could be a critical factor, but again we did not find any difference between the 23 TV programs that elicited retrieval of personal details and the other 27. Nevertheless the chance of getting any recall for these 23 TV programs was higher than it was for the two oldest TV programs broadcast in the late 50's: "Les Aventures de Mic" and "Télé-Pok", for which no report was collected and for which correct identification on the 4-FC was respectively 12.1% and 2.9%.

Interestingly, the median age of the participants at the time when these 23 TV programs were broadcast was 33 years old ($SD = 8.4$, range 18-52) which corresponds to a critical period in the developmental literature: adults above 35 years old are able to retrieve many autobiographical memories during this period of early adulthood (Rathbone et al., 2008; Rybash, 1999). This might have had an impact on the older participants' ability to recall the test TV programs in semantic memory.

This age effect was confirmed when considering participants' overall performance for the 50 test TV programs. Indeed we found that the best performers were the youngest subjects from the older group of participants.

To get access to the participants' memories of old TV programs, this study was carried-out in an ecological way in comparison to classical lab-based memory paradigms. However these real-life conditions lead to two main limitations. 1/ The population we tested was heavily biased towards females, with only three males among the 34 participants. This bias is largely explained by differences in participants' willingness to be involved in this study, together with a difference of mean life expectancy between the genders in France (French male: 80y,

French female: 85y). The small number of male participants made it impossible to find significant performance differences between the genders, but further studies would be useful.

2/ We don't know whether the 50 test TV programs were watched by all the participants tested. However we found that the earlier participants had a television in their house the better their overall performance was for the test videos.

The biggest question raised by this study concerns the nature of the underlying mechanisms of these very long-term memories. How is it possible that such dormant memory traces can survive for decades, even when there is no possibility of re-exposure to the original stimuli? So far, studies have mainly focused on the role of re-consolidation in the maintenance of long-term memory. Findings show that reactivations of long-term memories re-consolidate and strengthen memory traces and for episodic memories this occurs with the support of a more distributed ensemble of hippocampal–neocortical neurons (Moscovitch et al., 2006). Because the TV programs used had between 6 and 120 episodes, such re-consolidations might have occurred during the initial period of broadcast. At the same time, the retrieval of similar episodic events during the original broadcasting would allow the transformation of episodic information into semantic representations corresponding to a gist or schema created from all the episodes watched for the same TV program (Winocur and Moscovitch, 2011; Winocur et al., 2010). For one TV program, an episodic memory of each TV episode would then coexist with a generic semantic representation of all its episodes. According to this Trace Transformation Theory that follows the Multiple Trace Theory, episodic memories would rely on both neocortical and hippocampal structures whereas the semantic representation would be specific to the neocortex.

In the absence of reactivations after the original broadcasting, the synaptic weights of the neurons supporting these memories -in the hippocampus and/or in the neocortex- would not have been much reinforced. And so far, little is known concerning the maintenance of these dormant memories.

Our paradigm used audiovisual clips to reactivate these dormant memories and it is quite possible that such stimuli could be more efficient than static and unimodal stimuli (Furman et al., 2007). We strongly believe that the use of dynamic and multimodal stimuli should be more widespread and might provide valuable assistance in the diagnosis of memory loss or impairment in conditions such as Alzheimer's disease. We are open to providing the audiovisual clips to researchers interested in using the material.

It is as yet unclear how such long-term memories can be retained over decades and it should be noted that synaptic plasticity may not be the only possibility. For example, there is

evidence that some forms of memory can be transmitted epigenetically (Carone et al., 2010; Crick, 1984; Dias and Ressler, 2014) although it seems unlikely that such mechanisms could be involved in memories for old TV themes. We hope that further investigations will be carried out to understand the “life” of these dormant memories.

5. Conclusion

Compared to classical memory experiments the results we got might sound ridiculous. To me, they were by far beyond what I could expect when we started the investigation. About 40 years later, I had the proof that some individuals could retrieve spontaneously the title of a TV program with no more support than a 7s audiovisual excerpt. Even if this happened to only two participants out of the 34 we tested, it makes it clear that dormant memory traces can be recalled years later.

With Sophie we spent several months to gather all of this data. Between the lotto game and the snack at 4 pm, it was hard to convince our potential participants to perform a one-hour test in front of a computer. Especially men seemed to prefer playing pétanque in the backyard. It might sound cliché, but it was the reality we faced. Importantly, we were able to identify 23 TV programs that elicited some recall. Using only these 23 TV programs, we might be able to test more participants and make it more attractive to our target population. Indeed, with Simon and Nadège we thought of it as a first step before carrying it at a larger scale. So far, the data collected are promising but they might be more robust if hundreds of people would be tested.

6. Acknowledgements

This research was supported by the European Union’s Seventh Framework Programme (FP7/2007-2013)/ ERC grant agreement n°323711 to S.J.T. We thank the Institut National de l’Audiovisuel (INA) for giving us the opportunity to access French audiovisual archives and especially Hélène Bettembourg and Chantal Delmont. We thank the two anonymous reviewers for their constructive feedback that allowed major improvement in the writing of the manuscript.

7. Supplementary information

Participant	Gender	Age when tested	Year of TV acquisition	Hours of TV/day
1	female	76	1956	3.5
2	female	84	1960	2
3	female	81	1970	5
4	female	70	1958	1
5	female	90	1959	2
6	female	85	1963	6
7	female	91	1958	6
8	female	88	1967	3
9	female	84	1958	3
10	female	75	1963	4
11	female	76	1962	1
12	male	72	1956	3
13	female	78	1965	4
14	female	77	1963	0.5
15	female	75	1962	1
16	female	83	1960	9
17	female	78	1957	3
18	female	79	1963	2
19	female	52	1964	2
20	female	83	1970	5.25
21	female	73	1965	2.5
22	male	77	1963	3
23	female	71	1968	4
24	female	92	1966	1
25	male	74	1964	3
26	feminine	76	1970	2
27	female	74	1966	7.5
28	female	79	1968	6
29	female	84	1965	8
30	female	76	1964	5
31	female	76	1967	3
32	female	73	1965	2
33	female	89	1967	2.5
34	female	84	1970	4

Table 3.S1 Older participants' detailed information: gender, age, year of TV acquisition and daily number of hours of watching TV at the time of the recall.

Title	Date 1st episode	Date last episode	Rebroadcast	Number of episodes	Episodes duration (min)	Number of seasons
Famous TV programs						
30 millions d'amis	6 January 1976	On the air the 1 January 2015	No	> 1000	25	39
Age tendre et tête de bois	30 May 1961	1 April 2014	No	?	?	?
Agence tous risques	1 July 1984	8 March 1987	Yes	98	45	5
Alerte à Malibu	2 January 1991	1 January 2002	Yes	243	45	11
Bonne nuit les petits	10 December 1962	1 December 1973	Yes	568	5	5
Chapeau melon et bottes de cuir	7 January 1961	21 May 1969	Yes	161	52	6
Dallas	2 April 1978	3 May 1991	Yes	357	45	14
Des chiffres et des lettres	4 January 1972	On the air the 1 January 2015	No	?	?	?
Eurovision	24 May 1956	On the air the 1 January 2015	No	58	180	?
Fort boyard	7 July 1990	On the air the 1 January 2015	No	277	110	25
Inspecteur Gadget	24 October 1983	1 January 1986	Yes	86	22	2
Intervilles	17 July 1962	On the air the 1 January 2015	No	250	120	24
La petite maison dans la prairie	30 March 1974	21 March 1983	Yes	205	45	9
Mac Gyver	29 September 1985	21 May 1992	Yes	139	45	7
Magnum	11 December 1980	1 May 1988	Yes	162	50	8
Mission impossible	17 September 1966	30 March 1973	Yes	171	50	7
Motus	25 June 1990	On the air the 1 January 2015	No	> 1000	30	24
Stade2	28 December 1975	On the air the 1 January 2015	No	> 1000	70	39
Starsky et Hutch	30 April 1975	15 May 1979	Yes	93	48	4
Thalassa	27 September 1975	On the air the 1 January 2015	No	> 1000	60	37
Thierry la fronde	3 November 1963	27 March 1966	No	52	25	4
Wonder Woman	7 November 1975	11 September 1979	Yes	59	48	3
Test TV programs						
A dossiers ouverts	18 February 1974	22 March 1974	No	25	13	1
Acilion et sa bande	3 July 1978	31 August 1980	No	?	?	2

Alexandre Bis	25 July 1974	5 September 1974	No	6	78	1
Alice où es-tu?	16 August 1969	8 September 1969	No	20	13	1
Animal Parade	14 February 1972	25 February 1972	No	12	7	1
Bayard	16 January 1964	9 April 1964	No	13	25	1
Candice ce n'est pas sérieux	9 September 1969	1 October 1969	No	20	13	1
Champions	21 May 1964	12 July 1966	No	?	?	?
Commandant X	30 October 1962	17 July 1965	No	10	90	1
Courte Echelle	8 February 1974	8 July 1974	No	~100	150	1
En direct avec	3 October 1966	8 April 1968	No	?	?	?
Encore un carreau de cassé	11 aout 1960	29 April 1961	No	?	?	?
Fachoda ou la mission Marchand	X March 1977	1 April 1977	No	6	60	1
Faire Face	10 June 1960	9 February 1962	No	11	30	2
Frédéric le gardian	23 September 1965	16 October 1965	No	24	14	1
L'arche de Samsong	8 November 1972	26 Decembrer 1972	No	37	15	1
L'as et la virgule	19 November 1964	23 December 1964	No	8	45	?
L'Enjeu	X September 1978	13 June 1988	No	?	?	?
L'homme de l'ombre	29 July 1968	6 September 1968	No	30	13	1
La Prunelle	10 September 1968	3 December 1968	No	13	30	1
La boîte à malice	1 July 1979	30 August 1979	No	?	?	?
La vérité sur l'espionnage	4 December 1967	1 January 1969	No	13	50	?
La vie commence à minuit	28 July 1967	6 August 1967	No	20	13	1
Le blanc et le noir	14 January 1975	9 September 1975	No	?	?	?
Le comte Yoster a bien l'honneur	30 July 1969	5 October 1974	No	40	25	2
Le dernier petit ramoneur	6 October 1961	1 January 1961	No	?	?	?
Le Pèlerinage	5 April 1975	5 May 1975	No	24	13	1
Le roi qui vient du sud	8 February 1979	15 March 1979	No	6	52	1
Le Tourniquet	1971	1971	No	?	?	?
Le train bleu s'arrête 13 fois	8 October 1965	11 March 1966	No	90	13	1

Les Atomistes	12 February 1968	18 March 1968	No	26	13	1
Les Survivants	20 September 1965	6 October 1965	No	13	13	1
Les Aventures de Mic	2 April 1959	14 September 1960	No	52	15	1
Les complices de l'aube	1 September 1965	9 September 1965	No	9	15	1
Les demoiselles de Suresnes	15 April 1968	21 May 1968	No	26	13	1
Les dernières volontés de Richard Lagrange	4 July 1972	7 August 1972	No	30	13	1
Les diamants de palinos	31 December 1964	14 January 1965	No	13	13	1
Les facéties du sapeur Camember	20 September 1965	18 November 1965	No	50	5	1
Les sept de l'escalier 15	3 January 1967	6 February 1967	No	25	13	1
Malican père et fils	15 July 1967	27 July 1967	No	13	26	1
Miss	26 July 1979	16 August 1979	No	6	55	1
Objectif Demain	19 March 1979	16 december 1981	No	?	?	?
Poker d'as	31 October 1973	7 December 1973	No	26	13	1
Quand on est deux	18 November 1962	31 January 1964	No	26	13	1
Que ferait donc Faber ?	3 July 1969	21 August 1969	No	8	55	1
Télé-Pok	11 September 1957	11 February 1958	No	?	?	?
Teuf Teuf	30 October 1968	7 March 1969	No	~111	?	1
Un grand amour de Balzac	19 April 1973	31 May 1973	No	7	50	?
Vacances animées	14 March 1974	1 January 1977	No	120	45	3
Vol 272	3 May 1964	26 July 1964	No	13	26	1

Table 3.S2 Details of famous and test TV programs: broadcast period, number of episodes, duration of episodes and number of seasons.

For some TV programs the number of episodes was not known and an approximate numbers (~...) was calculated based on the days of broadcast and the dates of beginning and end of broadcast. Question marks represent missing data.

CHAPTER 4. Long-term memory and familiarity after 12 years

Abstract

In 2006 Mitchell demonstrated that implicit memory was robust to decay by showing that the identification of fragments of pictures seen 17 years before was significantly higher than for new stimuli. Is this true only for implicit memory? In this study, we tested whether explicit memory was still possible for drawings ($n = 144$) that had been presented once or three times, two seconds each time on average, about 12 years before. Surprisingly, our data reveal using two different but congruent experiments that a third of our participants (8 out of 24) were able to recognize pictures above chance levels. Preserved memory was mainly observed in young subjects, for stimuli seen 3 times. Despite the fact that confidence judgments were low, reports suggest that recognition could be based on a strong sense of familiarity. These data extends Mitchell's findings and show that explicit memory can also be robust to decay.

Presentation chapter

The data collected in the first experiment gave me some confidence. I saw with my own eyes the two participants who recalled spontaneously the titles of the TV programs. It was to me and to Sophie that participants reported their subjective impressions and memories about the TV programs. Yet, it took me some time to really believe in them, until I experienced it myself.

I remember this day when I was in a conference room for a meeting afternoon organized by the doctoral school of my university. Researchers were invited to speak about sensory information, and how we perceive it. After interesting discussions about the spectrum of light and synesthesia, the last speaker started his talk on a movie. Before he started, I remember saying to a friend that I did not know this movie. But during the first minutes of his introduction it was like his words ringed me a bell. Details about the movies sounded familiar to me and I started to think that it was probably due to similarities with a book that I read or with a documentary that I could have watched on the same topic. This familiar flavor that I got from the general introduction all made sense when the researcher showed a still snapshot from a particular scene of the movie. I already watched this movie! And as soon as I realized it, it was like being sent back in a specific time and space of my life. I was in my small apartment in Paris, at my desk that was also my dining room table. I was at that particular place, close to the wall that separated the leaving/bed room from the small kitchen. And I just remember this particular moment when I could see a light ray running from the top left of the screen to the keyboard of my laptop. In front of this natural lighting I could see all the dust on my computer, and especially around the on and off button. It was probably the sunset, at about 8pm and I know that it was time for diner. I could not tell more. But I had no doubt, what I was watching at this time was the movie: “The Taste of Cherry” a title that I completely forgot about. But I did not forget the faces of the characters, either the particular scenes of the movies, or the overall story... This dormant memory was still pretty young, 2 or 3 years old but it gave me the inner feeling that we were in the right direction.

From the first experiment I could tell that individuals could maintain the memory of TV programs for several decades in an inactive state. Some TV programs were broadcast not even a dozen times others more than a hundred and I would have thought that the number of broadcasts might have an influence on individuals' recall. However, I could not find any correlation between participants' performance and the number of episodes of a TV program, or with the whole exposure duration (duration of one episode time the number of episodes). As discussed in chapter 2, repetition might be the key for storing long-term memories. During the watching of a TV series, many repetitions occur. As an example, main characters can be seen hundreds of times just in one episode and even by watching a TV episode for a couple of minutes we might be able to store the repeated patterns in long-term memory. Therefore, TV series might not be the best to tackle the issue of the number of presentations.

My curiosity pushed me to go for a new project that was even more audacious than the first one. I remember when I heard through Nadège that Emmanuel was interested in testing participants' remote memory for very simple drawings. These drawings were seen no more than six times and about 10 years ago during a neuropsychological protocol in Marseille. I had the chance to conduct this research with the support of Eve who carried out the original experiment in Marseille, Sophie and of course Simon who was always looking for new challenges!

Unlike my first experimental design, this one was by far more classical. We knew that hundreds of drawings were seen by a group of participants during a specific protocol. We just had to contact them again and test their memory. For each participant I could provide precise figures such as the percentage of retrieval.

This time, the stakes were high. Assuming that memories might not be gone, I had to reactivate memory traces of drawings that were the most basic pieces of art that one could think of: Microsoft cliparts from 2001. Unlike TV programs, these drawings were not part of a specific associational network where semantic and contextual information might have been potential access doors. And worse, the status of drawings reduced them to frozen shapes on a white page. In such extreme conditions, I knew that I should force participants to try to retrieve the memory of these drawing even if they said not to have any trace left. For that I had to make it simple: every time two drawings will be presented, one they saw and one new. Participants would just have to tell which one they already saw.

Of course, we had to check that there was nothing different between the new drawings that we added in and the ones that were shown in the original experiment. Djamaa helped me gathering the new drawings and ran pilot controls. Everything was then ready for Sophie and Eve to test our participants.

1. Introduction

Implicit memory does not involve explicit or conscious recollection and the effect of memory is inferred from behavioral performance (Roediger, 2003). In 2006, Mitchell (Mitchell, 2006) published a research report demonstrating that implicit memory could be preserved on the long-term. In his experiment, participants were better at identifying fragments of black-and-white line drawings they had seen 17 years before than new drawings, even when they could not remember having participated in the original study. Even today, this finding holds the record for the duration of implicit memories, raising the question of the neurobiological mechanisms underpinning such ability.

In contrast, in explicit memory tests, participants are directly involved in the retrieval of previously learned items. This holds for recognition memory tasks where participants have to determine if an item has already been experienced or not. It is generally assumed that recognition in such tasks relies on two processes: familiarity, which is the ability to identify that an item has previously been encountered, and recollection, which implies the retrieval of contextual details associated with the item (Mandler, 1980). Some studies have suggested that explicit memory can extend to a few decades if not for a lifetime. Such studies have focused, for example, on the memory of old classmates (Bahrick et al., 1975) or on knowledge learned in school (Conway et al., 1991). However, all these studies used ecological material for which the duration of acquisition -reflected through the number of exposures and the duration time- was far longer than in Mitchell's experiment. Indeed, priming lasting for 17 years was found for drawings that had been presented for 1 to 3 seconds and at most 3 times. Furthermore, the material used in the explicit memory tasks was probably rehearsed either purposefully (knowledge learned at school) or simply by interest (recurring discussions about old classmates with friends, etc.). Therefore, it is unknown how long explicit memory can be maintained if exposure to the material is brief and if no rehearsal has taken place.

Recent neuronal models based on Spike Time Dependent Plasticity (Masquelier and Thorpe, 2007) show that a few exposures to simple stimuli can be enough to create hyper-selective neurons. Such neurons could potentially remain silent and their synaptic weights preserved until the same stimulus or a close representation would be presented again (Barron et al., 2016). According to this view, implicit but also explicit memories might be maintained for a life-time (Thorpe, 2011). We therefore hypothesized that subjects exposed to simple stimuli may be able to recognize at least some of them several years later, even with no exposure in the meantime.

In the present paper, we tested participants' explicit memory ($n = 24$, 14 female) for simple colored drawings ($n = 144$) that had been presented once or three times for roughly 2 seconds each time on average, from 8 to 14 years before. Because the stimuli were simple and not particularly interesting (a third of them were actually abstract figures difficult to verbalize), we thought that it was unlikely that the participants ever thought about them between encoding and test about a decade later.

2. Method

2.1 Participants

From 2002 to 2008 (mean year = 2004, $SD = 2.2$), 243 healthy control subjects participated in a neuropsychological protocol in the Timone University Hospital in Marseille (France). This protocol included four cognitive tasks among which a visual recognition memory test that we used as the basis of the current study.

We could contact 63 of these subjects, of whom 25 agreed to participate in a second experiment in 2016. One of them was discarded because of macular degeneration. None of the 24 other participants reported the presence of important neurological, psychological or psychiatric troubles since 2002 and are referred to as test participants in this paper (14 females, mean age = 60.8y, $SD = 17.5$, range: 34-84y). All subjects had normal or corrected-to-normal vision and audition. The overall cognitive abilities of participants beyond age 60 were assessed in 2002-2008 using the Mini-Mental State Examination (MMSE) (Folstein et al., 1975) from the French GRECO consensual version (mean MMSE score 29.5 out of 30, $SD = 0.7$, range: 28-30). All participants were tested again with the MMSE in the context of the current study with no deficit found for any of the participants (mean score = 29.5 out of 30, $SD = 0.9$, range: 27 – 30).

A control group of 24 naïve subjects (14 females, $M = 62.5$, $SD = 20.7$, range: 33-94y), who had never seen any of the drawings from the first experiment also took part in this study.

The current study was approved by the INSERM Ethical Evaluation Committee (CEEI, N°15-263) and all participants gave informed written consent before experimentation.

2.2 Initial test: the DMS-48

The basis of the current study is a visual recognition memory test, the DMS-48, that is widely used in memory clinics in French-speaking countries and has been described extensively before (see for example Barbeau et al., 2004; Didic et al., 2013). The DMS-48 is freely available for research purpose (<http://cerco.ups-tlse.fr/~barbeau/dms48.html>).

During an incidental encoding phase (participants were not told that it was a memory test), 48 pictures were first presented one after the other. For each stimulus, participants were simply instructed to tell if the picture was made of more or less than 3 colors. Two forced-choice recognition memory phases took place after 3 and 60 minutes. Different neuropsychological tasks were presented during the interfering phases. Participants were asked to identify the picture previously seen among two presented simultaneously (one old and one distractor each time). There were three conditions (examples in Figure 4.1). In the ‘Unique’ conditions, both old and distractors pictures were concrete objects with names and shapes (e.g. a rooster and a digger). In the ‘Paired’ condition old and distractors items had the same name and similar shapes and colors (e.g. two different snowmen). The ‘Abstract’ condition involved abstract patterns that were difficult to verbalize. A separate set of 48 new distractors was used for each of the test phase. All stimuli were presented on A4 white sheets of papers. Performance on the DMS-48 is expressed in percentage (maximum performance 100%).

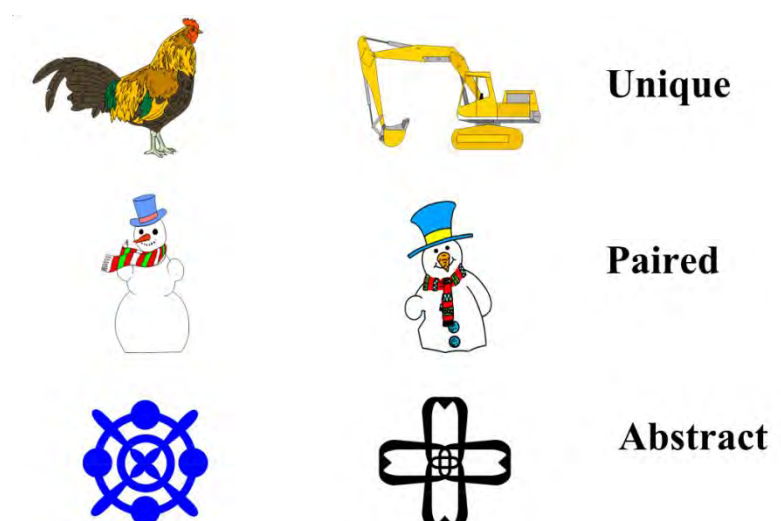


Figure 4.1 Examples of pairs of stimuli used in the DMS-48 for the three conditions “Unique”, “Paired” and “Abstract”.

2.2.1 Initial number of presentations of the stimuli

Overall, the 48 old stimuli were seen 3 times (during the incident encoding phase and the recognition phases at 3 and 60 minutes), whereas the other 96 stimuli that served as distractors during the recognition phases at 3 and 60 minutes were seen only once. This was the case for 19 participants. However, 2 subjects (participant 4 aged 64 and participant 5 aged

67) underwent a third recognition phase after 6 weeks in 2007 and were thus exposed 4 times to the old stimuli. The remaining 3 subjects (participant 2 aged 83, participant 3 aged 84 and participant 6 aged 63) underwent the whole protocol again 18 months later and were thus exposed 6 times to the old stimuli and twice to the distractors.

2.2.2 Initial time spent on the stimuli

The time spent to perform the incidental encoding of the 48 stimuli and to complete each of the 2 recognition tests was calculated for 19 out of the 24 participants. For 5 participants the time to perform the 2 recognition tests was not measured.

Participants spent on average 212 s (SD = 42.5, range: 150 - 327s) to perform the incidental learning task, 166 s (SD = 52, range: 83 - 310s) to complete the first recognition test and 159 s (SD = 47.4, range: 76-279s) to finish the second recognition test. By dividing the duration of these 3 experimental runs by the number of stimuli presented, we were able to get an estimate of the exploration time for each of the 144 pictures. On average, the 48 stimuli presented 3 times were seen for 7.8 s (SD = 1.7, range: 5-12.9s), whereas the 96 stimuli presented once were seen for 1.7 s (SD = 0.5, range: 0.8-3.2s).

2.2.3 Initial results

Participants' performance was at ceiling (after 3 minutes: mean = 98.4%, SD = 2.0, range: 92 – 100% and after 60 minutes, mean = 98.5%, SD = 1.9, range: 92 – 100%) thus, demonstrating that essentially all the targets had been fully processed and memorized at least for an hour.

2.3 Current tasks

In the present study, we used two successive recognition memory tasks to assess participants' ability to recognize the stimuli seen between 8 and 14 years earlier: a forced-choice task as close as possible to the initial one, but using a new set of distractor stimuli, and a Yes/No task to assess whether the findings could be replicated in a different context, i.e., were robust to experimental changes (Figure 4.2).

Because this is the first study of its kind as far as we know, we were careful in trying to keep environmental variables the same from the initial exposure to the present experiment, thus trying to respect the encoding specificity principle (Tulving and Thomson, 1973). We used a similar way to present the stimuli (printed on A4 white sheets presented in binders) and subjects were seen in the same hospital as initially.

The forced-choice task was based on 144 pairs of coloured drawings. Each pair was made of an old stimulus seen in the initial task 8 to 14 years earlier and a new one. Old stimuli were either the 48 targets seen several times or the 96 distractors seen only once in the initial test. The new items were chosen with respect to the original combination of categories ‘Unique’, ‘Paired’, and ‘Abstract’ ($n = 48$ for each). For each trial, the two items were displayed on the right and left-hand side of the sheet (counterbalanced). The pairing of the new items and the ones seen initially was shuffled across participants for the categories ‘Unique’ and ‘Abstract’. However, because it would have been difficult to find enough new pictures that would match the items of the ‘Paired’ category presented in the DMS-48, the same “Paired” combination was presented for all participants. The order of the presentation of the pairs of the stimuli was randomized across participants. The forced-choice task was divided into 3 blocks of 48 pairs of pictures. Trials were pseudo-randomized to equalize within each block the number of stimuli seen 3 times and once as well as the number of stimuli belonging to each combination categories (‘Unique’, ‘Paired’, ‘Abstract’).

For each trial, participants were invited to identify the picture that was presented in the initial experiment several years before.

This task was also presented to the 24 naive control subjects. Of course, they had not seen any of the stimuli before, so they were asked to say which one they “thought” had previously been presented. The reason for this procedure was that the new items used in the forced choice were selected in 2016. There could thus have been a risk that these new items were inadvertently systematically different from the previous ones, which would have introduced a bias in the experiment.

After completing the forced-choice task, the participants were invited to respond on a Yes/No recognition memory task. 144 out of the 288 stimuli shown in the forced-choice were individually presented, centred on A4-printed white sheets. These stimuli were equalized across the 3 categories ‘Unique’, ‘Paired’, ‘Abstract’ and included the 48 stimuli seen 3 times in the initial test, 48 stimuli seen once in the initial test and 48 stimuli from the new set of stimuli added to perform the forced-choice task of the current experiment. Only 48 stimuli were used for each condition to limit the duration of the Yes/No task. Thus, all stimuli used in the yes/no task had been presented in the forced-choice task.

For each stimulus, participants were asked to say whether the picture had been shown to them in the initial experiment (a “yes” response was expected) or if it was a new picture just seen in the previous forced-choice task (a “no” response was expected in this case).

Two participants refused to perform this task. One because she said she saw too many similar pictures in her life (she used to work with children in a day nursery) and the other one because it was too hard for him to differentiate drawings seen in the initial test from those seen previously.

2.3.1 Confidence responses and verbal reports

For each trial in the forced-choice and the Yes/No tasks participants were asked to rate the confidence of their response on a 5-point scale: 0: 'guess', 1: 'not sure at all', 2: 'not very sure', 3: 'fairly sure' and 4: 'sure'. Importantly, the participants were also invited to report the reason for their choice verbally if they used a confidence score of 3 or 4. As high confidence responses can be based on recollection, we wanted to assess whether participants would spontaneously recollect elements of the initial test that would help them recognize some stimuli.

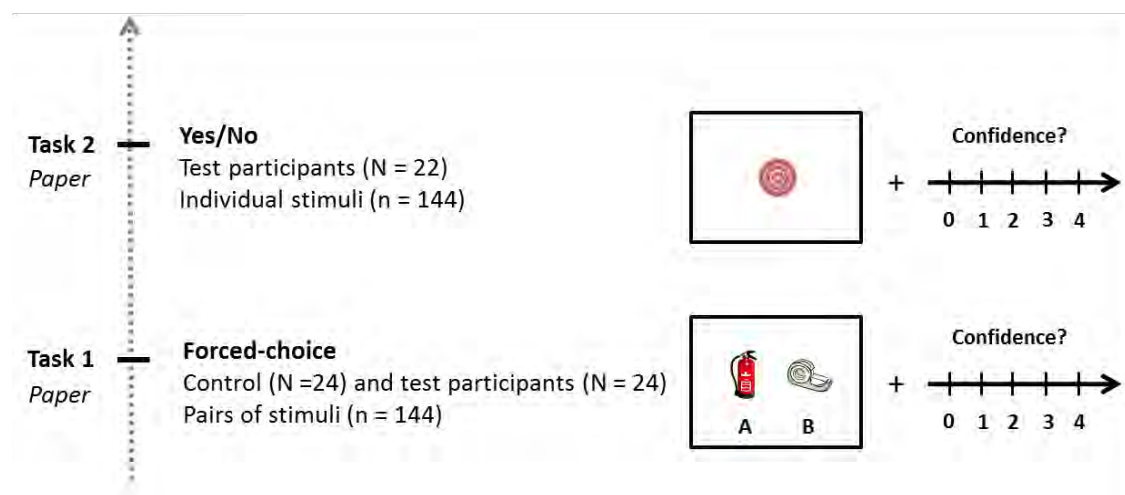


Figure 4.2 *The current experimental procedure consisted in two explicit memory tests: a forced-choice (task 1) and a Yes/No recognition (task 2).*

In the forced-choice, test participants (N = 24) and control participants (N = 24) were presented with 144 pairs of stimuli and participants were asked to identify which of the two stimuli was presented in the first experiment (DMS-48 performed in 2002-2008). In the Yes/No task the participants (N = 22) were asked to evaluate if the 144 individual stimuli corresponded to 'old' (seen in the DMS-48 about 10 years ago) or 'new' items (seen only in the previous forced-choice). In both tasks the stimuli were presented on A4-paper sheets and each response was followed by a confidence score rated on a 5-point scale.

2.3.2 *Performance and presentation of the results*

Performance in the forced-choice task was expressed as a percentage (maximum performance 100%).

To take into account the ratio imbalance between the “Yes” and “No” responses in the Yes/No task, performance was calculated using a sensitivity index: $d\text{-prime} = Z(\text{Hit rate}) - Z(\text{False Alarm rate})$ with Z being a z-score. Within each analysis, Z -score values were corrected (following Snodgrass & Corwin, 1988) for every participant if some of them had infinite scores. A d -prime value of 0 corresponds to chance-level performance and larger d -prime values correspond to better performance. As previously done for correct responses in the forced-choice task, a continuous representation of the data was used.

To get a better apprehension of the data, the estimated number of participants for a given correct percentage was also calculated using a kernel density estimation (Allen et al., 2012). This allowed us to get a continuous representation of the data that cannot be visualized via classical bar or box plots.

2.3.3 *Statistical analyses*

We followed the new statistics guidelines to test any potential effect in the data (Cumming, 2014). As recommended, we did not use null-hypothesis significance testing (NSHT) that relies on p-values. Instead we presented confidence intervals that offer a better understandability of the data and in our case of the chance level.

Chance level on the forced-choice was calculated based on the estimate of a binomial distribution (50%) associated with its 95% confidence interval (CI). This value gives an estimate of the chance level for one participant with an upper ($50\% + 95\%CI$) and a lower ($50\% - 95\%CI$) chance limit. As an example, for the 144 pairs of drawings presented in the forced-choice task, chance level ranged between 60 and 84 items ($144/2 \pm 95\%CI$). In this case, participants would perform better than the chance level if they could correctly identify more than 84 items. Repeating such measure across participants could lead to some uncertainty of this estimate at the group level. We considered that 95% of the participants should fall within the 95% confidence interval of the chance level. For 24 participants, this means that 22-23 participants should perform within the chance limits and only 1 or 2 participants should be out of the CI limits. In the case where more than two participants have performance beyond these chance limits, the population deviates from chance level.

The material and the data are freely available here:

<https://drive.google.com/drive/folders/0B1akb3XjfkkiyWHc3SW52QmEydlE?usp=sharing>

3. Results

3.1 Naïve control participants

The new items used in the forced choice task were selected for the purpose of the current study and could have been inadvertently systematically different from the stimuli chosen at the time the recognition memory task was designed, which would have introduced a bias in the experiment. To control for this, the performance of naive control subjects who had never seen any of the stimuli and could thus in principle not perform the task was compared to the chance level expected from a binomial distribution delimited by its 95% confidence interval limits. Control participants were on average 50.9% correct on the forced-choice task (SD = 5.9, range 41.7% - 63.2%) and were not different from the 50% chance level expected from a binomial distribution for 144 items ($CI_{95} = [41.6\% - 58.4\%]$) with only 2 participants performing above the upper limit.

3.2 Test participants' performance

In the forced-choice task, two drawings were simultaneously presented: one corresponding to a picture presented in the initial experiment about 12 years before, one corresponding to a new picture. Interestingly, the median distribution of test participants' performance was significantly different from the one obtained by the control group (55.2%, vs 51.4%, $U = 387$, $CI_{95} = [0 - 8.3]$, Mann-Whitney-Wilcoxon-test) showing a general upper shift of the whole population of test participants from chance level (Figure 4.3).

The probability to identify by chance the picture presented in the first experiment corresponded to the 50% of a binomial distribution, which varies according to the number of trials. The minimum and the maximum values around the 50% chance level corresponded respectively to the lower and upper 95% confidence intervals of the binomial distribution. Test participants' mean performance on the forced-choice was 55.3% (SD = 7.2, range: 40.3% - 69.4%) with 8 participants above the upper chance limit of 58.4%. This means that these 8 best participants could identify from 1 to 15 pictures (mean = 7.4, SD = 3.9) more than the maximum 84 pictures expected by chance ($144/2 + 95\% \text{ CI}$).

In addition, for the 22 test participants who performed the Yes/No task the d-prime was low (mean = 0.3, SD = 0.3, range: -0.5 to 1.1) but larger for the 8 best participants in the forced-choice task (mean = 0.5, SD = 0.3, range = 0.1- 1.1). Performance in the forced-choice and

Yes/No tasks was positively correlated ($r = 0.44$, $CI_{95} = [0.0 - 0.7]$, Pearson's correlation coefficient).

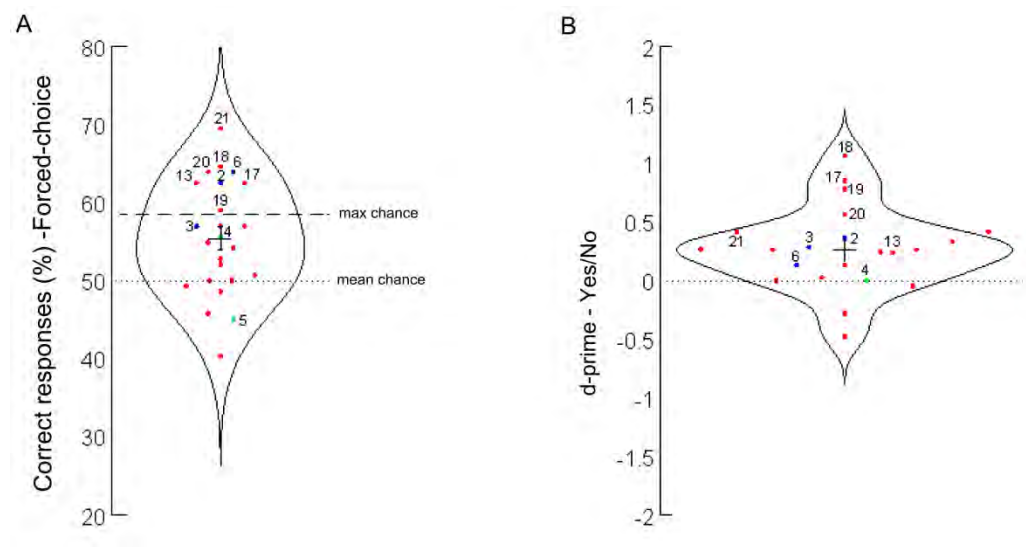


Figure 4.3 Test participants' performance on the current experiment.

(A) Performance on the forced-choice task ($N = 24$) for the 144 pairs of pictures. (B) Performance on the Yes/No task ($N = 22$) for 144 pictures. One dot is the performance of one participant. The upper chance limit in the forced-choice task is represented by the dashed line and the chance average by the dotted line in the forced-choice and in the Yes/No task. The number of the eight best participants in the forced-choice task is provided as well as the number of the participants who saw the old stimuli four times (green dots) or six times (blue dots). The estimation of the number of participants for a given correct level of response was also calculated to get a continuous representation of the performance. This kernel density estimation is represented by the surrounding shape (Allen et al., 2012). The cross represents the mean of the distribution.

3.2.1 Effect of the number of exposures

In the initial test performed between 2002 and 2008, 48 stimuli were seen 3 times (= old stimuli at the time) and 96 only once (= distractors at the time) except for 2 participants (participants 4 and 5) who were exposed 4 times to the old stimuli and 3 others (participants 2, 3 and 6) who were exposed 6 times to the old stimuli and 2 times to the distractors. Two of the subjects who had seen the stimuli more times than the others were among the 8 best subjects mentioned above (participants 2 and 6). In the current forced-choice task, the 24 participants were on average 57.4% correct ($SD = 11.0$, range: 37.5% - 72.9%) for the stimuli seen at least 3 times and 54.3% correct ($SD = 6.4$, range: 41.7% - 68.7%) for stimuli seen less

than 3 times (Figure 4.4A). Seven participants were above the upper chance limit of 64.8% for stimuli seen at least 3 times. These participants were able to correctly identify from 2 to 4 pictures more than the maximum 31 pictures expected by chance level in the forced-choice task. Four of them corresponded to participants who saw the stimuli only 3 times and 3 of them to the participants who were exposed 6 times to the stimuli. In comparison, only 2 participants were above the upper chance limit of 60.4% for stimuli seen less than 3 times. In the Yes/No task, participants' performance was higher for stimuli seen at least 3 times with an average d-prime of 0.4 (SD = 0.4, range: -0.4 – 1.0) compared to stimuli seen less than 3 times (n = 48) with an averaged d-prime of 0.1 (SD = 0.4, range: -0.6 – 1.1) (paired t-test $t(21) = 6.3$, CI_{95} difference for d-prime seen 3 times vs d-prime seen once = [0.2 – 0.4]) (Figure 4.4B).

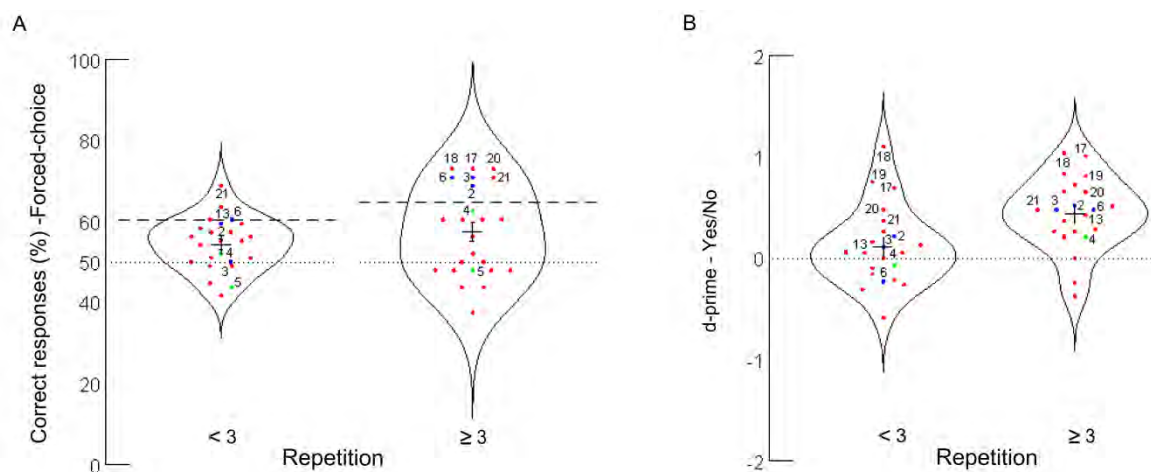


Figure 4.4 Participants' performance on the forced-choice and Yes/No tasks according to the number of times stimuli were seen in the DMS-48 (2002-2008): less than three times (< 3) or 3 times and more (≥ 3).

A. The distribution of the correct percentages in the forced-choice is presented with their respective kernel density estimation to get a continuous representation of the data. The upper chance limit is represented by the dashed line. The number of participants above the upper limit is displayed as well as the number who saw the old stimuli four times (green dots) or six times (blue dots) (1 dot = 1 participant). **B.** d- prime values for test participants in the Yes/No task are presented with their respective kernel density estimation. The eight best participants in the forced-choice task are shown as well as the number of the participants who saw the old stimuli four times (green dots) or six times (blue dots) (1 dot = 1 participant).

Chance average is shown by the dotted line in the forced-choice and in the Yes/No task and the cross represents the mean of each distribution in the two tasks.

3.2.2 *Stimulus type*

The stimuli used in this experiment belonged to three categories: ‘Unique’, ‘Paired’ and ‘Abstract’. No difference was found between these three categories for both the forced-choice and the Yes/No task.

3.2.3 *Other influencing factors*

We tested the potential influence of different general factors in the forced-choice task. We found that the performance of the participants was negatively correlated with their age ($r = -0.44$, $CI_{95} = [-0.7 \text{ to } 0.0]$, Pearson’s correlation coefficient), but not with gender ($r = -0.11$, $CI_{95} = [-0.49 \text{ to } 0.3]$, Pearson’s correlation coefficient) or educational level ($r = -0.0$, Pearson’s correlation coefficient).

3.2.4 *Confidence rating*

Every response was followed by a 5-point scale confidence level ranging from 0: ‘guess’ to 4: ‘sure’.

On average test participants were not sure of their responses for both the forced-choice (mean = 1, SD = 0.9, range: 0 - 2.4) and the Yes/No task (mean = 1.1, SD = 1.0, range: 0 – 3.2). In the forced-choice task, 46.8% (SD = 41.1) of the responses were rated as a “guess” (rate 0) and 23.3% (SD = 31.1) as “not sure at all” (rate 1). In the Yes/No task, 38.9% (SD = 42.3) of the responses were rated as a “guess” and 24.7% (SD = 24.7) as “not sure at all”.

Performance increased with confidence in both the forced-choice task (Figure 4.5) and the Yes/No task but this increase was not significant. The mean confidence of the 8 best participants was low in the forced-choice task (mean = 0.7, SD = 0.9) and in the Yes/No task (mean = 0.8, SD = 0.9) except for participant 2 who reported the highest score over the whole test participants: mean = 2.4 both in the forced-choice and the Yes/No task.

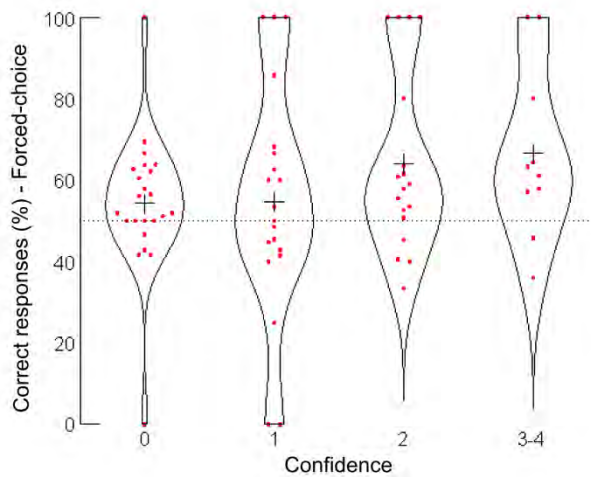


Figure 4.5 *Participants performance over the confidence level in the forced-choice task.* 0: “guess”, 1: “not sure at all”, 2: “a bit sure”, 3: “fairly sure” and 4: “sure”. The responses associated to “fairly sure” and “sure” responses were grouped together because of their small number. Chance average is shown by the dotted line and the crosshairs represent the means of the distributions.

3.2.5 Participants’ reports















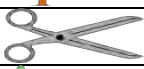





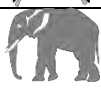







Participants were invited to provide a justification when they were ‘fairly sure’ (rate 3) or ‘sure’ (rate 4) of their responses.

Fairly sure responses were made 9.7% (SD = 16.8) of the time in the forced-choice and 10.3% (SD = 14.1) in the Yes/No task whereas sure responses were made 1.1% (SD = 3.2) of the time in the forced-choice task and 3.4% (SD = 10.1) in the Yes/No task (Table 1). Overall 61.7% (SD = 32.1, range: 19.7% - 100%) of the responses with a confidence of 3 or 4 were associated with some verbal reports in the forced-choice task and 25.8% (SD = 35.4, range: 0 - 100%) in the Yes/No task. Drawings associated with comments were correctly identified 71.4% (SD = 17.9, range: 46% - 100%) of the time in the forced-choice task and 77.3% of the time (SD = 32.6, range: 0 - 100%) in the Yes/No task.

Importantly, justifications reported were all about visual properties of the stimuli but did not make any reference to any contextual details (see Table 1 for examples from 3 representative participants for stimuli seen 3 times).

3.2.6 Participants' recollection of the initial test

Two participants out of 22 who were asked whether they recollected the initial test were not able to remember that they had performed it several years before. Interestingly, participant 5 was one of the eight best performers in the forced-choice task. None of the 22 participants reported having seen again or rehearsed any of the stimuli presented.

	Participant	3 times	New	Confidence	Report
UNIQUE	1			3	It speaks to me
	1			3	The drawing speaks more to me than the other one
	11			4	I can't remember having seen the sweater at all
	11			4	I remember having seen marine animals but not specifically the dolphin
	11			4	By elimination
	11			3	The shape of the building makes me think that I saw this drawing
	16			3	I think I saw it
PAIRED	1			3	I cannot tell why
	1			3	I do not think I saw anything symmetrical
	1			3	A matter of color
	11			3	Because of the originality of the drawing
	11			3	The shape seems more real
ABSTRACT	1			3	Because of the color
	11			3	By default because I cannot remember having seen the other picture



	16			3	It speaks well to me
--	----	---	---	---	----------------------

Table 4.1: Examples of reports in the forced-choice task for stimuli seen 3 times. The reports from 3 representative participants for which correct identification was associated with a confidence of 3 or 4 are presented according to the 3 categories: ‘Unique’, ‘Paired’ and ‘Abstract’.

4. Discussion

In this study we found that our group of participants was able to recognize simple colored pictures seen for a few seconds between 8 and 14 years before, above a group of naive participants who had not been exposed to the stimuli. Furthermore, a third of the participants (8 out of 24) performed above an estimate of chance that included an upper 95% confidence interval. Our best performer, who had been exposed to the pictures at most three times, was able to identify 15 pictures more than the 84 pictures expected by chance. Note that no instruction to specifically learn the stimuli was ever given to the subjects, even at initial encoding, which makes this performance even more remarkable. Mitchell’s findings (2006) revealed that the memory of stimuli seen briefly and not rehearsed could be maintained for more than a decade and retrieved implicitly. Our data provide strong evidence that memory can be retrieved explicitly. Indeed, subjects were able to choose between two propositions the picture they had seen before and endorsing it as old or new.

Importantly, an analysis of verbal reports provided for high confidence responses revealed that retrieval was never associated with the recollection of contextual details or any reference to the initial encoding stage. In contrast, subjects emphasized that details or feelings helped their recognition: they referred to the shape or to the color of the stimulus, the stimulus “spoke” to them, or they “thought” they had seen it. Hence, high confidence recognition of these items seems to rely on a sense of familiarity suggesting that familiarity may be used to recognize pictures seen only a few seconds even a decade after initial presentation. Although familiarity is supposed to be a mechanism used in long-term memory, it had never been demonstrated that it could be used so long after the initial encoding stage. Interestingly, some subjects also appeared to assess the novelty of the distractor in the forced-choice task, suggesting that correct recognition might have relied on a balance between familiarity and novelty assessment in some instances.

Most of the items (~ 85%) however were recognized with low confidence levels suggesting that participants were unsure exactly why they chose one picture over the other. In this case recognition might have relied on a fluency effect, the increased ease with which subjects process a picture due to repetition, which is in agreement with Mitchell' findings (2006). However, subjects were able to capture this signal and to attribute it to a previous exposure to make an explicit choice, even if unsure (Besson et al., 2015; Jacoby and Whitehouse, 1989; Voss et al., 2008). Even these responses thus appear to be based on familiarity.

The ability to maintain precise information about sensory stimuli has already been found in a couple of experiments where participants were able to discriminate pictures seen a few hours before from highly similar pictures (Brady et al., 2013). Other evidence indicate that memories can be maintained from a few days (Standing, 1973) to almost a year (Urgolites and Wood, 2013). They might constitute "episodic elements" that -if combined with contextual details- would help form an episodic memory (Conway, 2009). In this experiment we go even further and show strong evidence that perceptual memories can explicitly impact behavior for a decade.

However, this ability seems to depend on participants' age at encoding. Indeed, performance was negatively related to age. Interestingly, many memories from the early adulthood can still be recalled years later and is defined as a 'critical bump' in the developmental literature (Conway and Pleydell-Pearce, 2000). The critical bump is sometimes related to the important events that usually occur during this period of life such as first job, wedding, etc. Our results suggest that it could also be related to an intrinsic ability to form and maintain perceptual memories.

Importantly, we used a strict criterion to identify 8 participants with above chance level in the forced-choice task. However, we found also participants with performance around or even below 50%, demonstrating that explicit retrieval was not possible for them. The reason why some subjects perform above chance is unclear at present although age could be an important factor to take into account. It might be interesting to test the participants with poor performance on an implicit task to know whether some memory traces can still be captured. In this study we tested the recognition of simple colored-drawings. Evidence has shown that memory of pictures is better than words (Paivio and Csapo, 1973), a phenomenon known as the "picture superiority effect". Our results might only hold true for pictures. We used different types of visual stimuli (concrete or abstract) and different pairings of distractors. Surprisingly, our data do not show any strong difference between these categories, reinforcing

the idea that what we tested is mainly a pictorial memory rather than a combination of verbal or semantic and visual memory.

As opposed to priming, it has been shown that for explicit memory tests retrieval is influenced by the environmental context (McKone and French, 2001). Because we were unsure of what to expect, we were careful in trying to keep these environmental variables the same from the initial exposure to the present experiment, thus respecting the encoding specificity principle (Tulving and Thomson, 1973): subjects were tested in the same hospital and stimuli were presented on similar white sheets in similar binders using a similar forced-choice task. The impact of the respect of the encoding specificity on our results remains to be clarified.

Memory reactivations can occur during “on-line states” when an individual is voluntarily retrieving a specific memory (Nyberg et al., 2000) or during “off-line” states” when memories can emerge spontaneously during sleep or at rest (Diekelmann and Born, 2010; Peigneux et al., 2006; Wilson and McNaughton, 1994). “Off-line” reactivations play a critical role in the consolidation of a memory and occur within a few hours after learning (Gais et al., 2000; Girardeau et al., 2009). Although we can be confident that none the stimuli was ever seen between the initial encoding phase and the current experiment or that the subjects did not think about the initial encoding stage during this period, we could not control for any involuntary rehearsal that might have occurred during sleep or wakefulness (Rubin and Berntsen, 2009). However, “off-line” reactivations probably decrease exponentially over time (Frankland and Bontempi, 2005) and whether they still occur several years after encoding seems unlikely.

We show here that three presentations of a stimulus might be enough to create a familiar representation in very long-term memory, particularly in young subjects. The precise mechanisms allowing such feat remain to be clarified.

5. Acknowledgment

This research was supported by the European Union’s Seventh Framework Programme (FP7/2007-2013)/ ERC grant agreement n°323711 to S.J.T. We thank Jean-Michel Hupé for his constructive feedback in the statistical analysis of the data and Djamaa Atamena for her help in the preparation of the stimuli.

Discussion of the behavioral results

The results we got from two complete different experimental designs showed that long-term memories can be maintained in an inactive state. Collecting behavioral evidence was a first step that we initiated and that is far from being fully accomplished. As developed in chapter 1, every experimental design constrains participants' responses. Without such limitations it would have been impossible to make emerge specific experiences of retrieval. However, it is very likely that we might have missed interesting aspects during memory retrieval. In particular I think of all the episodic content of these memories. Indeed no episodic memory was retrieved in these two behavioral experiments. Is it because of the age of the participants? It is generally assumed that the retrieval of episodic information as well as the ability to predict the retrieval of episodic information is more difficult for older participants than younger ones (e.g. (Souchay et al., 2007)). To overcome such a deficit it is proposed to increase the amount of cues that would elicit the retrieval of a specific episode (Taconnat et al., 2007, 2004). This will be discussed at length in the general discussion.

A crucial point that remains to be understood is how is it biologically possible that simple pictures that have been presented just 3 times and for a few seconds only, are still able to drive explicit choices a decade later?

Indeed, in such extreme conditions, it seems unlikely that rehearsal of the stimuli could have played a key role in the preservation of the memory trace. The same question applies for old TV programs from which participants could not re-access since the acquisition period. These memories were in a dormant state (Dudai, 2004; Lewis, 1979) and they were again woken-up through the presentation of a specific stimulus such as a picture or an audiovisual clip. At the synaptic level, in the absence of reactivations after the original exposure, the synaptic weights of the neurons supporting these memories would not have been much reinforced. And so far, little is known concerning the maintenance of these dormant memories.

It is hard to conceive the preservation of these synaptic weights over time when considering a very large population of neurons as the support of the memory trace. In such networks the constant incoming of new inputs might interfere with previous memories that tend to be overwritten (Gardner, 1987; Hopfield, 1982). We suggest that the initial consolidation phase that enables the formation of a stable memory trace could be associated with the increased selectivity of a small number of neocortical neurons. This selectivity might rely on a simple STDP (Spike Time Dependent Plasticity) mechanism that allows simulated neurons to become selective to arbitrary input patterns if they occur repeatedly (Bichler et al., 2011;

Klampf and Maass, 2013; Masquelier and Thorpe, 2007) and even after a small number of repetitions (Andrillon et al., 2015). One of the characteristics of this sort of spike based learning is that synaptic weights are only modified if the target neuron fires (D'amour and Froemke, 2015; Markram et al., 1997). Selective neurons, that are not firing for any new incoming stimuli, would therefore be able to keep memories for a long time (Thorpe, 2011) in a silent way until their excitatory-inhibitory balance might be disrupted (Barron et al., 2016). It has been suggested that all of these inactive cortical neurons might form a kind of “dark matter” in the brain (Binzegger et al., 2004; Shoham et al., 2006; Thorpe, 2011). Re-exposure to a specific stimulus might elicit the firing of the otherwise inactive neurons which might trigger its recall and potentially the recollection of contextual specifics through binding in the hippocampus

Note that such STDP mechanisms do not rely on hippocampal circuitry and could occur in the neocortex. Interestingly, convergent evidence show that performance on the recognition memory task used in this study can be independent from the hippocampus (Barbeau et al., 2011) reinforcing the idea that these silent neurons could be localized in the neocortex using mechanisms independent of the memory system supported by the hippocampus. Re-exposure to a specific stimulus might elicit the firing of the otherwise inactive neurons which might trigger its recall and potentially the recollection of contextual specifics through binding in the hippocampus.

Presentation of the imaging experiments

Although positive and exciting, our behavioral results were tiny effects that could have been easily missed if our designs were not well focused or rigorous enough. Even before starting the experiment we used to hear participants saying that they were sorry for us because they might have forgotten everything! If they accepted to participate in our studies it was more for having a bit of fun and/or for being kind with young scientists, than to prove that they had good remote memories. Interestingly, for many participants this severe self-judgment tempered throughout the experiment. In the strict framework of the task, participants could themselves judge with more subtlety if they could retrieve the memory of a particular item or not.

Behavioral results are the end of a cascade of brain processes that unfold in space and time. I was therefore curious to investigate all of these stages together. To see them from closer, I had to record the brain when participants would be retrieving long-term memories. Although techniques evolve, it is still difficult to get the precise location and timing of specific neuronal activations at the same time. Because my particular focus was on the temporal dynamics of retrieval, I decided to get time-resolved signals that I could acquire easily using electroencephalography (EEG) techniques. And to my surprise, I realized that very few studies had been carried on this topic. It was a great opportunity for me to start my investigations! But given the subtlety of the effects, it was risky to rely only on classical EEG analyses such as Evoked Response Potentials (ERPs) and time-frequency analyses. Too coarse, they might fail to work. In addition, I decided to use decoding analyses that have already shown their power to detect small effects in a large variety of paradigms.

In chapter 5 I introduce EEG and I provide technical information about the analyses performed in chapters 6 and 7.

CHAPTER 5. Detecting the retrieval of long-term memories on a EEG signal

1. Introduction

While behavioral measurements can already tell us a lot about the underlying neuronal activity involved during a specific task, functional brain imaging techniques have been developed to get a direct access to the brain activity. These techniques are more or less invasive and allow spatial and/or temporal fine-grained resolution resulting in a variety of neural signals (see Figure 5.1 for a non-exhaustive list).

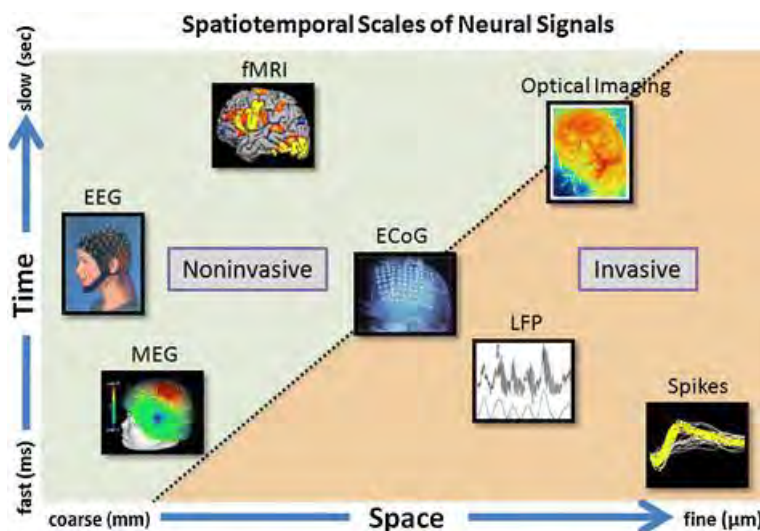


Figure 5.1 Spatiotemporal scale of neural signals for commonly used neuroimaging techniques.

Note that this list is non-exhaustive. Among the non-invasive techniques EEG and MEG have a very good temporal resolution whereas fMRI provides a good spatial resolution. Invasive techniques: ECoG and iEEG allow both good temporal and spatial resolution with especially iEEG that can record LFP and spikes generated by a very small numbers of neurons. From (Kim et al., 2015).

Non-invasive techniques are separated into two main classes: techniques that allow direct measurement of brain activity such as EEG (Electroencephalography) and MEG

(Magnetoencephalography) and techniques that indirectly measure neural activity such as fMRI (functional Magnetic Resonance Imaging).

2. EEG compared to other neuroimaging techniques

2.1 EEG

Discovered in 1929 by Hans Berger, EEG is the oldest neuroimaging technique which is still broadly used because of its simplicity. It consists in putting electrodes on the scalp of participants to record in real-time the electrical activity generated by neuronal depolarizations. Nowadays, electrodes are about 1cm of diameter and can be fixed on a cap that participants put directly on the head. An electrolytic gel (e.g. Ag/Cl) is applied to maximize the conduction between the scalp and the electrode. The number of electrodes present on an EEG cap would vary depending on the required spatial resolution. While a 256-electrode cap is commonly used to monitor patients sleep activity, a 64-electrode cap would be usually enough to cover the different scalp regions. Electrodes position follows international standards such as the 10/20 system that covers the occipital, parietal, temporal, central and frontal regions of the scalp (Figure 5.2).

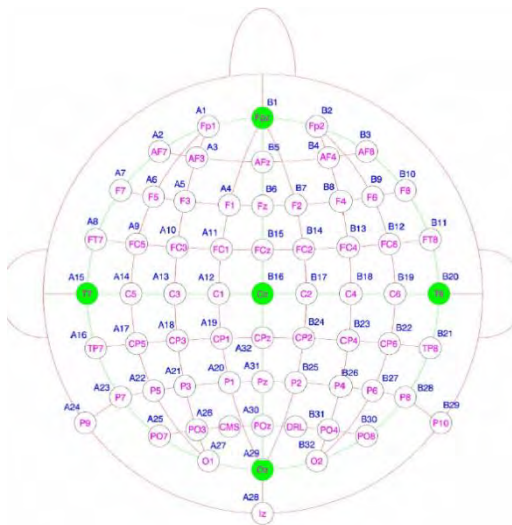


Figure 5.2 Lay-out of a 64-electrode cap using the BIOSEMI system.

At the top: frontal region and at the bottom: occipital region.

2.2 Comparison

The magnetic field produced by the electrical brain activity can also be recorded using MEG, a technique available since the 70's. The generated magnetic field being very weak (about 10-100 Femtotesla; 1 Femtotesla = 10^{-15} tesla) compared to the magnetic fields from the surrounding environment (about 10^{-8} tesla), specific (and expensive) requirements are

necessary to get access to the signal. Participants are recorded in a magnetically shielded room with very sensitive magnetometers: SQUIDs (superconducting quantum interference device) on top of the head.

Because the nature of the signal that is recorded by EEG and MEG is different: electrical or magnetic, the source of the brain activity captured is different. EEG records activity perpendicular to the scalp that is mostly generated by the pyramidal cells in cortical gyri and in the depth of the sulci. MEG records activity parallel to the scalp that is mostly generated by the pyramidal cells from superficial part of the sulci. EEG and MEG signals provide high temporal but poor spatial resolution.

Conversely, fMRI techniques have a high spatial but poor temporal resolution. Using the MRI technique which measures the hydrogen nuclei to reconstitute an image of the brain, fMRI allows the detection of changes of blood oxygenation (BOLD for Blood-Oxygen-Level-Dependent) in a specific region indicative of the neuronal activity in that region.

It is important to note that the signal acquired by M/EEG techniques contains very rich spatial information (Stokes et al., 2015). Indeed, it has been recently demonstrated that it is possible to differentiate orientations of specific gratings using MEG signal (Cichy et al., 2015). Considering that every different type of orientation is encoded in specific cortical columns in the visual cortex that are made of not more than a few hundred neurons each, this shows the extreme spatial precision of MEG and EEG. EEG and MEG are nevertheless considered as techniques of poor spatial resolution given the extreme difficulty to identify which specific region (source) generates the signal recorded at the scalp level. Indeed, source localization corresponds to an inverse problem for which there is never a unique solution...

Non-invasive techniques offer a lot of potential for research investigation and can even be coupled (such as combined EEG/fMRI) but remain limited to a certain time-space-scale. Higher precision can be provided using two invasive techniques: EcoG (electrocorticography) and intracranial electroencephalography (iEEG). It is important to note that such invasive techniques are used at first for medical purposes and in specific cases such as for patients with pharmacology intractable epilepsy. The EcoG consists in placing a grid of electrodes directly on the cerebral cortex. It covers a large and superficial area compared to the implanted electrodes in iEEG. Indeed intracranial electrodes used in iEEG can detect spikes and Local

Field Potential (LFP) generated by the activity of single neurons or small neuronal assemblies into deep brain structures.

The selection of a technique might depend on the scientific question we want to investigate.

For long-term retrieval we might want to ask:

- 1/ What brain regions are activated? and
- 2/ How does retrieval unfold over time?

The first question deals with spatial mapping and can be addressed using the fMRI technique whereas the second question can be more easily tackled via EEG. As we have presented it in the General introduction, very few data exist on the time-course of long-term retrieval. We therefore decided to investigate this field by using the EEG technique. In the following section we describe the techniques that we will use in the next two chapters.

3. Description of the techniques used

3.1 ERPs analyses

To compare two experimental conditions, the simplest way is to look at the ERPs that are small voltage (in μV) evoked by the brain in response to a specific stimulus. ERPs are already visible in the raw EEG signal and need little data processing. Because in principle, no transformation needs to be applied to the data, ERPs can provide a very precise estimate of the brain activity in the range of the millisecond. They are therefore of considerable interest when we want to study the time-course of the brain activity.

Along the history of ERP analyses, several main components were found to be consistent across many experiments such as the N100, P300.... A common way for naming these components consists in associating the letter P or N for respectively positive and negative peaks with their latency. Classical components corresponding to memory retrieval processes are the FN 400 for familiarity judgement and the LPC (Late Positive Component) for stimulus recollection (Curran, 2000). The FN 400 is a mid-frontal negativity occurring from 300 to 500 ms after stimulus onset and the LPC corresponds to a positive component between 500 and 800 ms after stimulus onset visible on parietal channels.

It is assumed that ERPs are both time-locked and phase-locked to an event, meaning that for a given time point we expect to get roughly the same voltage amplitude and phase across trials. For this reason, ERPs of a given condition are usually averaged across trials to get rid of the

noise present at a single-trial level. Because of this restriction, ERPs cannot capture information that is not time or phase locked to a particular event. Useful information might be therefore lost. Another issue is that it is hard to clearly know from which underlying neuronal activity the ERPs emerge. Indeed, little is known regarding the correspondence between neuronal activity and ERPs.

3.2 Time-Frequency analyses

While ERPs analyses are the best to provide temporal information they can miss valuable information and are difficult to interpret in terms of neurophysiological substrate. Time-frequency analyses have been proposed to give additional information and in particular about the contribution of specific oscillations in the EEG signal.

3.2.1 Brain oscillations

Brain oscillations are usually divided into five main wave categories or rhythms that have been linked to specific cognitive functions: the delta (0.5 - 3 Hz), theta (4-7 Hz), alpha (8-12 Hz), beta (13-30 Hz) and the gamma rhythm (30-100 Hz). In addition to these five main brain rhythms other types of oscillations can be found such as the ripples and fast-ripples that are oscillations of very high frequency generated in the hippocampus.

Frequencies in the theta and the gamma band are particularly interesting for studying memory retrieval. Indeed, theta waves are generated both in the hippocampus and in the neocortex and are strongly linked to memory encoding and retrieval (e.g.: (Hasselmo and Stern, 2014; Jacobs et al., 2006)) The role of gamma oscillations is less clear but evidence support their contribution to retrieval process (Düzel et al., 2003; Gruber and Tsivilis, 2008).

Interestingly oscillations are present in the brain at any spatial scale of investigation: from the recording of single neurons to the EEG signal. This fractal organization makes it easier to link oscillations recorded at the scalp level and the underlying neuronal mechanisms. The relationship between the two is far from being fully understood but increasingly advances have been made. As an example, a recent study showed that the restoration of gamma oscillations in mice enables the clearance of the A β -protein accumulated in the brain region, a protein that is known to be a clear factor of development of the Alzheimer disease (Iaccarino et al., 2016).

3.2.2 Time-Frequency analyses

The contribution of specific frequencies during a task can quickly change over time. Time-frequency analyses have been developed to track these differences.

- Basic example: Fourier Transform

To get the different frequencies that composed an EEG signal the basic idea is to apply a Fourier Transform to the signal. This corresponds to multiplying the signal with sine waves of different frequencies. Because this multiplication applies to vectors (EEG signal, sine wave) the multiplication is in fact a dot product where each element of the EEG signal is multiplied by the corresponding element of the sine wave. To get the involvement of each of these frequencies over time, the different sine waves are simply slid over time. This approach is called convolution and corresponds here to the dot product successively calculated at each time window using a given sin wave.

- Morlet wavelets

Sine waves are used in the Fourier Transform but cannot describe any changes over time because their amplitude always varies between positive and negative values without ever stabilizing. Other functions are therefore used to compute the convolution with the EEG signal and in particular Morlet wavelets that correspond to a sine wave windowed by a Gaussian (Cohen, 2014; Grossman and Morlet, 1984). The shape of this kernel provides a good trade-off between temporal and frequency resolution and was therefore used in the studies presented in the next two chapters. To prevent edge artifacts when the kernel is applied to a new time window, Morlet wavelets can be tapered using a specific function called window. Again specific windows exist and their use would vary depending on the precision wanted (Hanning, Hamming, Gaussian...). We decided to use the Hanning or Hann window which is an inverted cosine function that has for main advantage to taper fully the data to 0 at the beginning and at the end of a time-window (Cohen, 2014).

- Length of the time-window?

The length of the time window used to perform the analysis is critical to get a trade-off between frequency and time resolution. This length should be at least long enough to measure 1 cycle of a given frequency (which means 1sec for 1Hz) and preferentially 3 cycles of the given frequency meaning that the lowest frequency would be 3 Hz for a 1s-window. The

length of the time-window can then be reduced for higher frequencies to maintain a good resolution.

Since Ebbinghaus's first experiments in the late 19th century (Ebbinghaus, 1885) we know that in the absence of any rehearsal, there is a fast decrease of retrieval capacities in the days following learning. This decrease might come from a drop of neuronal activity that might be difficult to detect using the classical techniques that we have just presented above. To have more chance to dissociate the cerebral activity corresponding to a picture seen let's say, one month ago compared to a new picture, we decided to use multivariate pattern analyses (Haxby, 2012; Haxby et al., 2001).

3.3 MVPA for decoding subtle effects

3.3.1 MVPA for spatial and time resolved signals

- **MVPA for fMRI**

MVPA was first developed in fMRI studies. Over the years, the gain of spatial resolution of the fMRI resulted in the increase of number of voxels (pixels in 3D) in a specific region. Nowadays, a typical voxel size is about 2*2*2 mm. In classical fMRI analyses the voxels that respond significantly to a specific condition are selected and averaged together to reduce the noise in the data. However these two steps: voxel selection and spatial averaging lead to a reduction of the signal. Indeed, the information contained by the voxels that do not reach significance might be lost as well as the fine-grained spatial resolution of the pattern of voxels. To increase the power of the signal by taking into account the multidimensionality of the data -represented by multiple voxels in fMRI- multivariate pattern analyses were developed. Using MVPA all the pieces of information coming from individual voxels that contribute specifically to a condition are gathered. It allowed the detection of very detailed information about specific mental states such as edge orientation (e.g (Kamitani and Tong, 2005)).

- **MVPA for MEG and EEG data**

When considering time-resolved signal (such as MEG or EEG) the same multidimensional problem arises. Data multidimensionality comes from the high number of channels that record the same signal for a given time. Indeed several dimensions are contained in the cerebral activity: space (from multi-electrodes), time, frequency, phase and power. MVPA on time-

resolved signal was first applied for Brain Computing Interfaces (BCI) (for a review see: (Curran, 2003)) and its only since 2010 that decoding was used in cognitive tasks (Duncan et al., 2010; Schaefer et al., 2011). Following that, an increasingly number of studies demonstrated the power of MVPA for MEG and EEG data in the field of perception (Carlson et al., 2013; Cauchoix et al., 2016, 2014; Isik et al., 2014), memory (Fuentemilla et al., 2010; Jafarpour et al., 2014), audition (King et al., 2014) or even for taste discrimination (Crouzet et al., 2015).

3.3.2 MVPA: how does it work?

MVPA or also decoding is based on machine learning where the basic idea is to predict specific characteristics of new information. It is performed by an algorithm in two steps: training and testing. When the data used by the algorithm are individual trials rather than the average across several trials, it is called single-trial decoding analyses.

- Training

The algorithm is first trained on a subset of data so that it can learn specific characteristics of the data set. If the characteristics to be learnt were selected in advance and attributed to a specific type of signal (data were labeled) it is called supervised learning. In the case of unsupervised learning the signal was not labeled meaning that the algorithm has to find the relevant characteristics to be learnt.¹ It has been shown that repetition of similar patterns might be a characteristic that can be easily learned without any supervision (Masquelier and Thorpe, 2007). In the experiments developed in the next two chapters we decided to use supervised learning to train the algorithm on two categories (e.g ‘seen 3 weeks before’ and ‘new’).

- Testing

Once trained on a subset of data, the algorithm is tested on the remaining data subset. It is important to note the strict independence of the training and the testing data sets: if some data was used for training the algorithm it must not be used again for testing the algorithm. In addition, the same number of trials should be used for each category tested. The sample size of the training and testing sets varies according to the type of classification performed. In the experiments presented in chapters 6 and 7, 90% of the data were randomly used for the

¹ Note that this distinction is not always as simple and that semi-supervised learning exists as well.

training set and the remaining 10% of the data for the testing set. This random split is called Monte Carlo cross-validation and is performed multiple times with each time the generation of a new random training set (and its corresponding testing set). Monte-Carlo cross-validations give a larger range of possible combinations than k-fold cross validations where data are split into fixed and exclusive folds (Refaeilzadeh et al., 2009). In the case of 10-fold cross validations, the data are split into 10 fixed folds, with one fold being considered once as the testing set.

- Which algorithm to use?

Several types of algorithms can be used to compute decoding analyses. The ones preferentially used for time-resolved signals correspond to generalized linear models or Support Vector Machines (SVM). In generalized linear models the outputs can be discriminated by a line whereas SVM tries to find the hyperplane that best separates the data, where the hyperplane could involve multiple dimensions. In the next two chapters we performed a logistic regression that belongs to the class of generalized linear models, to get the probability that a given trial in the testing set belongs to one of the two categories trained on. A L2-regularization was applied to the logistic function to decrease the effect of very high or very low probability values. This penalization was performed by setting a Gaussian prior value to the regression coefficients. Note that L1-regularization exists as well and corresponds to the setting of a Laplacian prior value but does not give better results than the L2-penalty.

- Performance

Decoding performance is usually measured using the Area Under the ROC Curve calculated using the ratio between the Hits rate and the False Alarm rate. This sensitivity measure ranges from 0 to 1 where 0.5 is the chance level and 1 corresponds to the best performance.

When the same time point is used for training and testing the algorithm, the performance of the algorithm can be plotted for each testing time in a graph where the x-axis corresponds to the testing time and the y-axis to the decoding performance.

It is also possible to test the algorithm on other time points than the one used for training to see if the algorithm can generalize its predictions to the other time points. This creates a time generalization matrix where the y and x-axes now correspond to the training and testing times. In this representation, the diagonal of the matrix corresponds to the decoding with the same training and testing point. Unlike its simple diagonal, the matrix representation reveals

information concerning the temporal dynamic underlying the cognitive process. For example, a single square-shape pattern might reflect the activation of a single neuronal network whereas a diagonal-shape pattern might involve the rapid succession of several neuronal networks (J-R King and Dehaene, 2014).

4. Conclusion

We described three different methods to study the time-course of retrieval in long-term memory and adapted to EEG studies: ERPs, Time-Frequency and MVPA. This list is far from being exhaustive (or even optimal) but they might already provide some interesting information.

CHAPTER 6. Decoding long-term memories of short audiovisual clips

Abstract

Access to memories becomes harder as time elapses. Extensive literature exists on subsequent Remember/Know/NEW differences on brain activity but little is known about the direct impact of the retention interval length between the learning of new information and its recall. In this study, the EEG activity of 11 participants was recorded while snapshots from videos that had been watched 3 weeks ($n = 125$), 1 day ($n = 125$) or a few hours before ($n = 125$) were displayed among snapshots from videos never seen before ($n = 375$). Whereas ERPs and Time-Frequency analyses were only able to differentiate snapshots recently seen from new ones we showed that MVPA analyses could decode memories of snapshots seen 3 weeks before. As for snapshots seen a few hours or one day before, decoding matrices revealed a sustained pattern suggesting the involvement of a unique network of brain areas in the retrieval of audiovisual content. Importantly, we found that decoding scores were correlated with participants' retrieval performance.

Chapter presentation

In the previous chapter I discussed the interest of the EEG technique for analyzing the temporal dynamics of long-term retrieval. I presented three methods for measuring differences between experimental conditions: ERPs, time-frequency and decoding analyses. I had now to find my experimental conditions that I wanted to compare with. I was surprised to find that most of the EEG literature that designed memory paradigms used the contrast retrieved/forgotten. I could have done the same for long-term memories. But I was not very convinced that it might be the best contrast to study because I had no control for this subjective variable. I could have ended up with a strong imbalance between the number of retrieved and forgotten items making it impossible any comparison. Especially for decoding analyses I knew that I would need about a hundred of trials by condition. So to test a specific effect, I wanted to create a design that would directly control for this effect. To me this makes sense, and I was reassured when I read a paper of Rufin VanRullen that directly points out the “common fallacy” of analyzing EEG data based on subject reports (VanRullen, 2011). Therefore I decided to design an EEG experiment that was very similar to behavioral paradigms. My variable would be the time and more precisely the retention interval between the encoding and the testing phase. I could therefore compare the brain processes involved during long-term retrieval according to this variable.

Memory retrieval can be associated with a wide range of subjective experiences that strongly depend on the experimental design. Recall tasks are demanding because retrieval occurs spontaneously or from partial information (cued-recall). The completion of such process would take some time, a benefit when we want to study the whole cascade of brain processes involved in the retrieval of the items. Therefore, I decided to create a recall task where a cue would be presented at the beginning of the trials to elicit retrieval.

Most of the memory paradigms use words or pictures but I was not really convinced that they would be the best stimuli to study recall after a few weeks especially if hundreds of stimuli had to be learnt. To have more chance to be retrieved, I wanted to make sure that every stimulus that I would present would correspond to a distinct subjective experience. After having done my first experiment with TV programs I was sure that audiovisual clips would be the best. In addition, they were like short scenarios of everyday life events that could provide a better insight of what is recall in real-life conditions than simple words or pictures. And a great thing from using clips is that the cue to elicit recall would be just a still frame from the whole animation. I decided to collect short audiovisual clips from the Internet that were the

most diversified as possible. With Estelle's help I ended up with 750 of them, enough for me to test the same participants for three different retention intervals plus having a set of new stimuli. I decided to test this variable for each participant instead of having separate groups to make sure that performance would only be a result of the retention interval.

I remember when I said to Nadège that I would test participants for clips that they had to watch a few hours, one day or three weeks earlier, she looked at me like I was getting mad. How come? By showing them so many clips they won't be able to retrieve anything! This was her word: interference. Participants might get confused if they had to take in more than 100 clips in less than an hour. But I had in mind the experiment from Brady et al. (Brady et al., 2008) that Simon liked to talk about. In this study, participants were presented with 2500 pictures of real-world objects displayed for only 3 seconds. Surprisingly, six hours later, participants were about 90% correct in identifying the items previously seen from closely similar new items. How was it possible? The answer is that every picture corresponded to a specific basic level category, thus limiting potential interference between the items.

I was then sure that if I would present about 100 short audiovisual clips to my participants, they would be able to retrieve the information at least several hours later. To make sure that the memory trace would be maintained for longer delays I thought that it might be better to present the clips twice.

My job was then to try to get clips that would be as diversified as possible. Indeed, first, I tested some pilot participants and I could tell that after watching about 30 western clips participants got confused when they had to tell if the picture presented corresponded to a new western clip or one they already saw. But then, my set being more diversified, I got better behavioral results for a week of interval. Nadège seemed more relaxed and everything was ready for me to test and record participants' EEG activity until three weeks after having seen the first clips.

1. Introduction

As time goes on, memories fade away (Ebbinghaus, 1885). Successful recall can be the demonstration of a recollection process that retrieves contextual details about the memory or only based on familiarity judgment (Mandler, 2008, 1980; Yonelinas, 2002). In general, the ability to retrieve contextual details is high for recent memories but decreases quickly over time. After a few weeks, it is more likely that the retrieval of long-term memories relies on familiarity judgment or semantic knowledge rather than on recollection processes (Conway et al., 1997; Herbert and Burt, 2004; Piolino et al., 2009). Recollection and familiarity can be measured in a cued-recall task where partial information is provided to elicit memory retrieval (a cue). In such a task, the Remember/Know/New paradigm (R/K/N in short) is a typical procedure to assess participants' subjective recall. The R/K paradigm was originally developed to evaluate episodic knowledge and semantic knowledge (Tulving, 1985a) with respectively Remember and Know responses. This experimental procedure is now currently used to assess recollection (Remember) and familiarity (Know) although the equivalence familiarity = "Know" is not always true (Gardiner and Richardson-Klavehn, 2000; Williams et al., 2013). Several markers from electroencephalographic (EEG) recordings have been linked to familiarity and recollection processes. Specifically, studies showed that familiar and remember responses of visual stimuli (words or pictures) emerge from 2 different Evoked Response Potentials (ERPs) components. The first one, the FN400, a midfrontal negativity occurring within 300-500 ms after stimulus onset has been linked to familiarity. The second one is a late positive component (LPC) that kicks in between 500 and 800 ms in the left parietal region and has been linked to remembering ((Curran, 2000; Duarte et al., 2004; Düzel et al., 1997; Tsivilis et al., 2001) and for a review see (Rugg and Curran, 2007)).

The difference between familiarity and remembering is also reflected in terms of frequency bands, with an increase of power in the theta band (4-7 Hz) and in the gamma range (30-100 Hz) during respectively recollection and familiarity judgments (Gruber and Tsivilis, 2008). Despite the vast EEG literature on the topic of recollection and familiarity only a few of them directly tested the effect of retention interval between the learning and the recall of new material. ERPs studies showed that the 2 components persist over retention intervals ranging from several seconds (Nessler and Mecklinger, 2003) to several minutes and until one day (Curran and Friedman, 2004; Wolk et al., 2006). Interestingly, oscillations studies show an increase in gamma power with memory load (Düzel 2003, Gruber 2004, Howard 2004) suggesting a link between gamma frequency and long retention intervals (Jacobs et al., 2006). ERPs and Time-Frequency analyses are respectively classical univariate and bivariate

analyses that are usually performed on specific (or across) channels and across trials. This considerably reduces the multidimensional aspect of brain activity resulting in a decrease of overall information that can be partially overcome by Multivariate Pattern Analysis (Cauchoix et al., 2012). Indeed, it has been shown that MVPA improves signal-to-noise ratio and enable the detection of subtle details otherwise invisible (e.g (Kamitani and Tong, 2005; Nishimoto et al., 2011)). This technique has been recently used to study episodic memory using MEG data (Jafarpour et al., 2014, 2013) and is particularly interesting for dealing with the dynamics of the neuronal processes involved (J-R King and Dehaene, 2014). However, to the best of our knowledge, no EEG data exist concerning participants' recall of material acquired a few weeks before. In addition to short retention intervals: a few hours and one day, we decided to investigate participants' EEG activity ($n = 11$) during the recall of short audiovisual clips learnt 3 weeks before. Recall was elicited by a short presentation of snapshots of previously seen or new audiovisual clips and assessed by a R/K/N paradigm to track familiar or fully recollected items. We combined ERPs, Time frequency and MVPA analyses to study the brain dynamics involved during recall.

2. Materials and methods

2.1 Participants

Eleven healthy subjects (6 females, mean age = 24 y, SD = 1.9) gave written informed consent and participated to the experiment approved by the INSERM Ethical Evaluation Committee (CEEI).

2.2 Stimuli

Seven-second audiovisual clips ($n = 750$) downloaded from the internet were used in the experiment. Clips were selected to be diversified and unknown from participant, speechless and without lyrics. Half of them were shown to a group of participants ($N = 6$) during the learning sessions and the other half of the clips to the second group of participants ($N = 5$). For all audiovisual clips, one frame of the 7s-clip was chosen as a representative snapshot of the clip and presented during the recall phase. The same 750 snapshots were presented for all the participants: snapshots new for a group of participants corresponded to the ones seen by the other group of participants during the learning sessions.

2.3 Procedure

The experiment included two phases: an explicit learning of audiovisual clips and a recall test of these clips (Figure 6.1). The explicit learning phase was composed of 3 viewing sessions

occurring respectively 3 weeks (3W, mean interval = 21days 2h43min, SD = 16h54min), 1 day (1D, mean interval = 1day 2h21min, SD = 2h39min) and several hours (Hours, mean interval = 4h37min, SD = 1h06min) before the recall phase. In every session, 125 audiovisual clips (duration: 7 s, size: 320*240) were shown to participants (N = 11) on a computer screen. Every session was split in two blocks of 65 and 60 videos. All clips were repeated once inside a block but were not presented again in any other session blocks. To control for participants attention during viewing, subjects were asked to make a button-press response every time a clip was shown for the second time. Participants could also indicate that they already knew a clip by pressing a key assigned to this purpose during the clip presentation and the corresponding trials were discarded from further analysis. Overall, participants watched 375 audiovisual clips during the 3 learning sessions. Every clip was presented twice. The duration of a learning session was about 50 min. Participants were instructed that a recall test on the audiovisual clips will be performed 3 weeks after the first learning session. The testing phase performed 3 weeks after the first learning session was a cued-recall task of 750 trials and during which participants EEG activity was recorded. Each trial started with a fixation cross of varying duration (600-1000 ms) followed by either a snapshot of a clip seen during the learning sessions (n = 375) or a new one (NEW) (n = 375) and presented for 400 ms. After picture offset, participants were asked to press as fast and as accurate as possible, with the same hand, one of the 3 keys (randomized between participants) corresponding to R, K, N judgements. Before testing, participants were instructed about the meaning of those 3 response possibilities as follows: ‘Remember’: you identify the picture as previously seen in one of the 3 learning sessions and you can replay the related video in your mind by anticipating the following scenes of the short scenario. ‘Know’: you identify the picture as previously seen in one of the 3 learning sessions but you can’t retrieve to which video it belongs to. ‘New’: the picture does not belong to any of the videos shown during the 3 learning sessions. Participants were then invited to judge their response on a 5-point confidence scale (1: “Not sure at all”; 2: “A bit sure”; 3: “Fairly sure”; 4: “Very sure”; 5: “Completely sure”). The duration of the recall phase was about 1h15. Learning and recall phases were programmed using the Psychtoolbox, a Matlab toolbox.

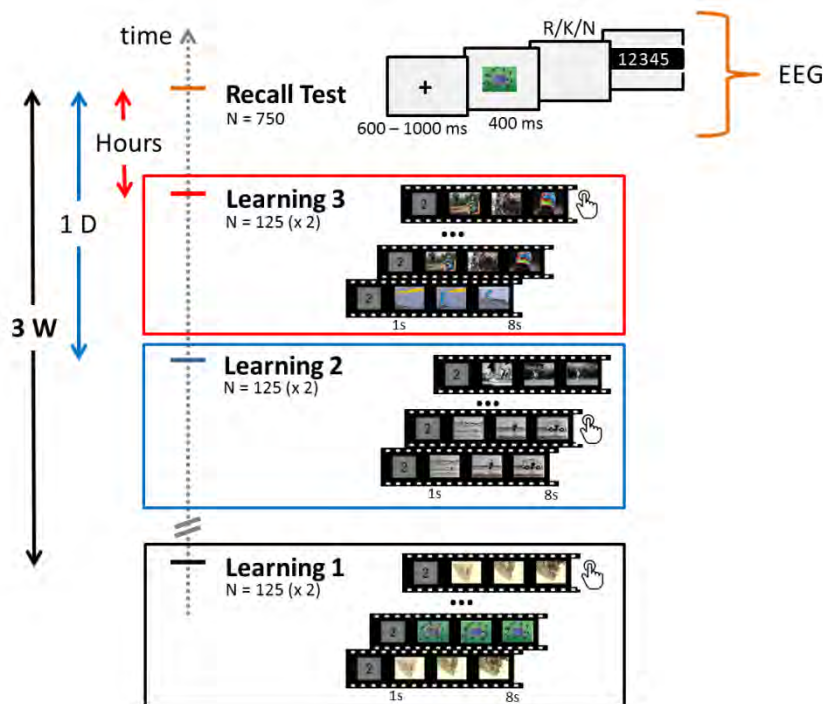


Figure 6.1 Experimental design.

Every participant watched 375 different videos during 3 learning sessions: 3 weeks (Learning 1), 1 day (Learning 2) and a few hours (Learning 3) before the recall test. For each learning session 125 short audio-video clips (7s) were shown twice on screen and participants had to make a button-press response every time a video was presented for the second time. Memory test was a cued-recall in which snapshots of clips seen during the learning sessions ($n = 375$) or new ones ($n = 375$) were briefly presented on screen (400 ms). After picture offset, participants were invited to make as fast and as accurate as possible a R/K/N judgement on the picture and rate the confidence on their response on a 5-point scale. EEG was only recorded during the recall phase.

2.4 EEG acquisition and analysis

All the 11 participants were included in the EEG analysis.

2.4.1 EEG acquisition and preprocessing

During the recall phase, participants' brain activity was recorded with a 64-channel cap connected to a BioSemi EEG amplifier (5 kHz) and with a sampling frequency of 1024 Hz. Data were filtered with a finite impulse response Blackman band-pass filter with a lower cutoff frequency of 0.1 Hz to avoid filtering artifacts (Acunzo et al., 2012) and a higher cutoff

frequency of 100 Hz with a transition bandwidth = 1 and then down sampled to 256 Hz. Trials were then epoched from -1s to 2 s relatively to picture onset, their baseline [-0.2 s 0] removed and re-referenced based to average reference. Artefacts were removed by performing ICA. For all participants, trials with extreme values for a channel ($>100 \mu\text{V}$ or $< -100 \mu\text{V}$) or abnormal trend (slope max: $75 \mu\text{V}$, $R^2 > 0.3$) or abnormal distribution (kurtosis outside 5 standard deviations of the kurtosis for each single electrode) were excluded from the analysis after visual inspection. All the preprocessing steps were done with EEGLab (Delorme and Makeig, 2004).

2.4.2 ERPs analysis

Trials were split according to the four retention intervals: HOURS, 1D, 3W and NEW for snapshots respectively seen a few hours, 1 day and 3 weeks before or never seen in the learning sessions or according to participants subjective responses: REMEMBER, KNOW and NEW responses. Differential activity was measured on 2 by 2 analysis by contrasting one condition with a NEW condition (retention interval or subjective response) and statistical difference was performed using Student t-test with a statistical threshold of $p < 0.05$ after False Discovery Rate (FDR) to correct for multiple comparisons.

2.4.3 Time-frequency analysis

Time-frequency analysis on every channel across participants was performed on the whole epoch from -1s to 2s by using a 3-cycles wavelet with a Hanning-tapered window for the lowest frequencies and increased half as fast as number of cycles. The window: 256 samples/1s wide was applied 200 times at an average step size of 2.57 samples (10.04 ms) and the results were oversampled 16 times. Significance level was calculated by using 50 bootstrap windows computed on the baseline ($p\text{-value} < 0.05$). A mask was then applied to restrict significant results to values that were significant for every participant and across at least 20 channels.

2.4.4 Decoding

Multivariate pattern analyses were conducted on the same data used for the ERPs analysis. The classification was performed on a 2 by 2 analysis with one condition constituted by trials of NEW snapshots and the second condition containing snapshots seen during a specific learning session (HOURS, 1D or 3W). Trials were randomly split 120 times according to the Monte-Carlo cross-validation procedure: 120 Cross Validations (CV). For each cross-

validation (CV), the number of trials was equalized between conditions 90% of the trials were used by the algorithm as a training set, the remaining 10% as a test set. Signal was normalized for the training and test sets. A linear classifier (L2 –regularized logistic regression, (Fan et al., 2008)) was trained on individual trials at each time point from -200 to 800 ms and tested on individual trials on the same and all the other time points (temporal generalization). The performance of the classifier for every participant and at each test time point was evaluated by using the Area Under the Curve (AUC). The average decoding performance across participants resulted in a matrix decoding where the x- and y-axes respectively represented training and testing times. Chance level decoding was calculated by performing the same classification on randomly permuted labels. Statistical significance was reached when the average across the 11 participants was higher than chance level for a duration of at least 30 ms of the training or testing time (paired t-test, $p < 0.05$, FDR corrected). Decoding analyses were performed by using an open source script developed by Sebastien Crouzet that can be downloaded at <https://github.com/scrouzet/classifyEEG> and that was used in (Crouzet et al., 2015) and in (Cauchoux et al., 2016).

3. Results

On average, eight out of 375 videos (SD = 9.8) were considered already known by participants and for each subject, the corresponding trials were discarded from behavioral and EEG analysis.

3.1 Behavioral results

3.1.1 Learning phase

Three learning sessions were done respectively 3 weeks (3W), one day (1D) and a few hours (Hours) before the recall phase. To make sure that participants were attending the videos, they were asked to make a button-press response whenever a video was presented for the second time. The percentage of HITs (correct button press) and FAs (wrong button press) were almost at ceiling and was stable over the different learning sessions: 3W: HITs = 96.7% (SD = 3.2, range: 90.2% - 100%) and FAs = 0.3% (SD = 0.4, range: 0-0.8%), 1D: 96.3% (SD = 4.7, range: 86.5% - 100%) and FAs = 0.5% (SD = 0.6, range: 0-1.7%), HOURS: 95.9% (SD = 5.5, range: 83.1% - 100%) and FAs = 0.2% (SD = 0.4, range: 0-0.8%). These results show that the encoding of the videos was similar across the 3 learning sessions (one-way ANOVA for HITs: $F(2,30) = 0.1$, CI_{95} -Difference (Learning 1 vs Learning 2) = [-4.4 5.3], CI_{95} -

Difference (Learning 1 vs Learning 3) = [-4.0 5.8], CI_{95} -Difference (Learning 2 vs Learning 3) = [-4.4 5.3]).

Inside a learning session, videos were presented twice inside a block of 65 or 60 videos. On average, the median number of videos between the two presentations of the same video was 30.4 (SD = 0.7, range: 30-32) and varied between 0.5 (SD = 0.5, range: 0-1) and 112.4 (SD = 3.4, range: 110-121) for each participant. We found that participants missed a repeat when the median number of videos between the first and the second presentation was 48 videos (SD = 22.0, range: 9.5-75.5) whereas for repeats that were correctly identified this number was 30.2 (SD = 0.7, range: 29-31). Because the number of misses was very small compared to the number of Hits (participants were almost at ceiling), we could not test whether the number of videos between a repeat significantly affect participants' performance.

3.1.2 Recall phase

During the recall phase, snapshots from videos seen by the participant ($n = 375$) or new pictures ($n = 375$) were briefly presented on screen. Participants were asked to make a R/K/N judgment as fast and as accurate as possible and to rate the confidence on their response on a 5-point scale.

- Performance over the retention interval

Participants were 90.06% (SD = 9.5), 85.1% (SD = 10.3) and 64.5% (SD = 18.1) correct in identifying snapshots seen respectively HOURS, 1D and 3W before and were 79.7% (SD = 13.3) correct for to identify NEW pictures. The percentage of correct responses for snapshots seen 3W before was significantly different from the performance for snapshots seen HOURS (CI_{95} -Difference = [10.4 - 40.7]) and 1D before (CI_{95} -Difference = [5.4 - 37.7]) (one-way ANOVA: $F(3,40) = 7.68$) (Figure 6.2A).

Participants' confidence average was 4.7 (SD = 0.3), 4.6 (SD = 0.4) and 3.9 (SD = 0.8) for the conditions HOURS, 1 D and 3W respectively and 3.5 (SD = 0.9) for NEW snapshots. The condition HOURS was significantly different from the conditions 3W (CI_{95} -Difference = [0.0 - 1.5]) and NEW (CI_{95} -Difference = [0.4 - 1.9]) (one-way ANOVA: $F(3,40) = 7.64$). It shows that participants were not only able to consciously retrieve the information but they also had a conscious access to the probability of being right or wrong. Subjective report was strongly negatively correlated ($r = -0.65$, Pearson's correlation coefficient) with the speed of the recall. Median reaction times increased over retention intervals: 1.04 s (SD = 0.12), 1.14 s (SD = 0.16) and 1.53 s (SD = 0.34) for snapshots correctly identified and respectively seen in the

conditions HOURS, 1D and 3W with the condition 3W significantly different from the conditions HOURS ($CI_{95}\text{-Difference} = [0.23 - 0.74]$) and 1D ($CI_{95}\text{-Difference} = [0.13 - 0.64]$) and was at 1.27 s (SD = 0.20) for NEW snapshots (one-way ANOVA: $F(3,40) = 9.5$) (Figure 6.2B) .

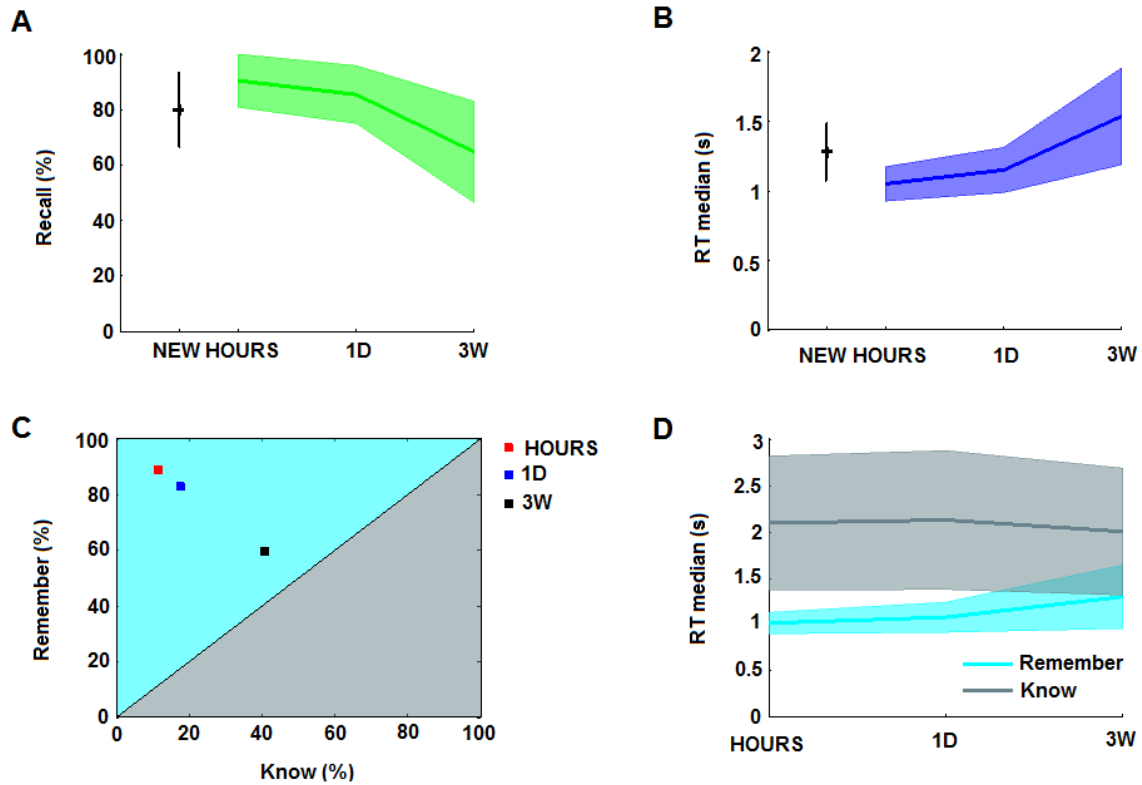


Figure 6.2 Behavioral performance: HITs responses on average across participants.

A. Percentage of HITs over the retention interval: snapshots seen hours, 1 day or 3 weeks before (mean and standard deviation). **B.** Reaction times (s) for HITs over the retention intervals (median and standard deviation). **C.** Repartition of Remember and Know for HITs responses over the 3 retention intervals (%). **D.** Reaction times (s) for correct Remember and correct Know responses over the 3 retention intervals (median and standard deviation).

- Remember, Know, NEW analysis

Overall participants' subjective responses were distributed as follows: Remember = 33.9% (SD = 11.4), Know = 15.8% (SD = 7.4) and New = 50.3% (SD = 11.8). This shows that for snapshots previously seen, participants reported being able to recollect the related video than simply knowing it (paired t-test: $t(10) = 3.9$, $p < 0.01$, $CI_{95}\text{-Difference} = [7.9 - 28.3]$). Recall performance was high for Remember and New responses with respectively 93.0% (SD = 6.0) and 81.4% (SD = 7.4) of correct responses and significantly lower for Know responses: 55.3% (SD = 20.4) (one-way ANOVA: $F(2,30) = 24.34$, $CI_{95}\text{-Difference} = [24.0 - 51.3]$, $CI_{95}\text{-}$

Difference = [12.5 - 39.8]). The average confidence was significantly the highest for Remember responses: 4.7 (SD = 0.2), and lower for NEW: 3.4 (SD = 0.9) and Know responses: 3.0 (SD = 0.5) (one-way ANOVA: $F(2,30) = 23.19$, $CI_{95}\text{-Difference} = [1.1 - 2.4]$, $CI_{95}\text{-Difference} = [0.6 - 1.9]$, $CI_{95}\text{-Difference} = [-1.1 - 0.2]$).

The ratio of correct ‘Remember’ over correct ‘Know’ responses decreased over retention intervals: 87.4% (SD = 15.7), 81.4% (SD = 16.4), 55.6% (SD = 23.31) for snapshots respectively seen in the conditions HOURS, 1D and 3W with the condition 3W significantly different from the conditions HOURS ($CI_{95}\text{-Difference} = [12.0 - 51.6]$) and 1D ($CI_{95}\text{-Difference} = [5.9 - 45.5]$), (one-way ANOVA: $F(30,2) = 8.85$). This is consistent with the literature showing a decrease in the ability to recollect past information for retention interval of weeks and more (Conway et al., 1997; Herbert and Burt, 2004; Piolino et al., 2009) (Figure 6.2C).

Reaction times for correct Remember responses was on average 1.01 s (SD = 0.1), 1.07s (SD = 0.2) and 1.30 s (SD = 0.3) for snapshots seen in the conditions HOURS, 1D and 3W respectively (Figure 6.2D) with a significant difference between the conditions HOURS and 3W ($CI_{95}\text{-Difference} = [0.04 - 0.52]$; one-way ANOVA: $F(2,30) = 4.69$). These objective measures were negatively correlated with subjective reports ($r = -0.58$). By contrast correct ‘Know’ responses were stable over retention intervals: 2.08 s (SD = 0.7), 2.11 s (SD = 0.7) and 1.99 s (SD = 0.7) for the conditions HOURS, 1D or 3W respectively ($F(2,29) = 0.09$). Overall, median reaction time was slower for KNOW (mean = 2.00s, SD = 0.7) compared to REMEMBER (mean = 1.07 s, SD = 0.1, $CI_{95}\text{-Difference} = [0.49 - 1.37]$) and NEW responses (mean = 1.27s, SD = 0.2, $CI_{95}\text{-Difference} = [0.29 - 1.17]$) (one-way ANOVA: $F(2,30) = 15.17$). Slower reaction times for KNOW responses are due to our experimental design that pushes participants to make preferentially fast ‘Remember’ responses. Late KNOW responses simply indicate that they were done when recollection failed after about 2 s.

To summarize, behavioral results indicate that participants were explicitly retrieving information from memory either on a recollection mode or based on familiarity judgment. Participants’ performance showed that recall directly depends on the time since videos were seen. To better understand how recall is performed and varies according to the retention interval, EEG analyses were carried on during the recall phase. Our investigation used classical uni and bivariate analyses (ERPs and Time-frequency) as well as multivariate patterns analyses (MVPA).

3.2 EEG analysis

3.2.1 Differential analysis, ERPs and time-frequency

Trials corresponding to the 3 retention intervals: HOURS, 1D and 3W were compared with trials of NEW pictures. ERPs analysis was performed on the mean average across participants for $t = -200-800$ ms after picture onset. Significant channels were found for a threshold at $p < 0.05$ (paired t-test, FDR corrected). The ERP topography (Figure 6.3A) shows that significant difference for HOURS/NEW appeared at $t = 500$ ms after picture onset in a first central channel (C3) and became stronger at $t = 600$ ms after picture onset mostly in the central and parietal electrodes. This differential activity remained stable until the end of the epoch ($t = 800$ ms). For 1D/NEW, significant channels were found at $t = 500$ ms post-stimulus onset and involved a distributed network of fronto-centro-parietal sites. Overall, until the end of the epoch, the same topography applied for 1D/NEW as for HOURS/NEW but with less significant channels. This topography was representative of the contrast Remember/NEW performed independently of the retention interval (Supplementary information, Figure 6.S1A) that might be a marker of the recollection of audiovisual clips. For the contrast 3W/NEW no significant channels were found in the whole epoch.

Time-frequency analysis on the band-pass signal (0.1-100 Hz) revealed significant positive difference only in the theta band. Difference was found for the early retention intervals: HOURS and 1D but not for the condition 3W when compared respectively to NEW pictures (Figure 6.3B). More precisely, for the condition HOURS/NEW a significant positive difference (bootstrap $p < 0.05$, mask applied) emerged in the slow theta band (~ 3 Hz) from the picture offset ($t = 400$ ms) to the end of the epoch ($t = 800$ ms) and in the fast theta band (~ 6 Hz). For the condition 1D/NEW a significant positive difference (bootstrap $p < 0.05$, mask applied) was also found in the slow theta band from 400 to 800 ms post-stimulus onset but not in the fast theta range. Interestingly, the combination of a positive difference in the slow and fast theta range is strongly present when contrasting all the REMEMBER responses with NEW responses (bootstrap $p < 0.05$, mask applied) (Figure 6.S1B) and might be representative of the recollection of the audiovisual clips. By contrast, when collapsing all the KNOW with the NEW responses a strong positive difference (bootstrap $p < 0.05$, mask applied) in the 3-5 Hz frequency range was present from about 200 to 400 ms post-stimulus onset. This component did not show up on the analysis performed on the retention intervals probably canceled out by the significant negative difference (bootstrap $p < 0.05$, mask applied) for the same time-frequency points in the case of REMEMBER responses.

Overall, by using classical uni- (ERPs) or bivariate (Time-Frequency) analyses we demonstrated that the processing of snapshots seen at most a day before revealed to be significantly different from NEW pictures. Difference was mainly captured on the late ERPs component: $t = 500-800$ ms by involving the theta band (3-7 Hz) in particular and might correspond to the recollection of the audiovisual clip as shown by the analysis performed on REMEMBER/NEW responses. However, these analyses failed to capture significant differences for longer retention intervals whereas 3 weeks after learning participants' performance show some recall. We decided to use MultiVariate Pattern Analyses on our EEG data to test whether snapshots seen 3 weeks can be differentiated from NEW snapshots.

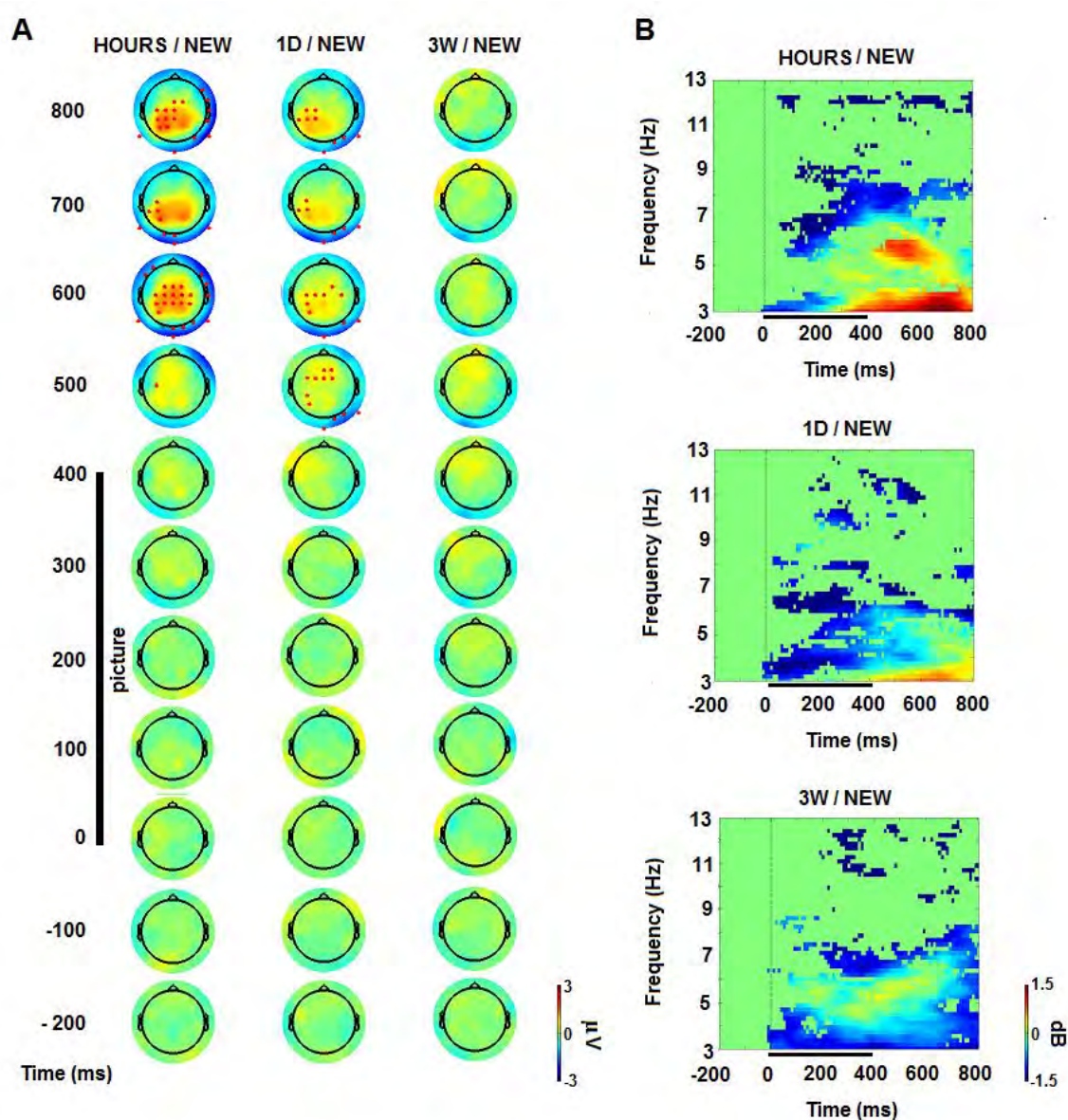


Figure 6.3 Differential EEG activity between snapshots seen in the 3 learning sessions (HOURS, 1D and 3W before the recall) and NEW snapshots.

Figures represent the average across the 11 participants for $t = -200-800$ ms post-stimulus onset. A. ERPs topography. Red dots correspond to significant channels ($p < 0.05$, FDR corrected). B. Time-frequency analysis. Only significant values (bootstrap $p < 0.05$, mask applied for 20 channels, all participants) are colored; non-significant values are plotted in green.

3.2.2 Early ERPs

The early time-course of memory recall on average across participants was analyzed before snapshot offset ($t = -200-400$ ms) for every channel and for 2 conditions (t-test, FDR corrected, threshold at $p < 0.05$) (Figure 6.4). When comparing the conditions HOURS versus NEW, the earliest difference was found for channel TP8 at $t = 100$ ms post-stimulus onset followed by a significant difference on the fronto-parietal sites: FP2 and FPz at respectively $t = 110$ and $t = 120$ ms post-stimulus onset. Difference were also found for the centro-parietal channel CP1 at $t = 120$ ms and $t = 140$ ms and CP3 at $t = 160$ ms post-stimulus onset. When contrasting the conditions 1D versus NEW, early activity was found for the channel C1 at $t = 280$ ms, then in the channel PO3 at $t = 290$ ms and in the frontal sites: AF8 at $t = 290$ ms, F7 at $t = 370$ ms and $t = 390$ ms. No significant difference for the condition 3W vs NEW were found in the ERPs neither before or after stimulus offset.

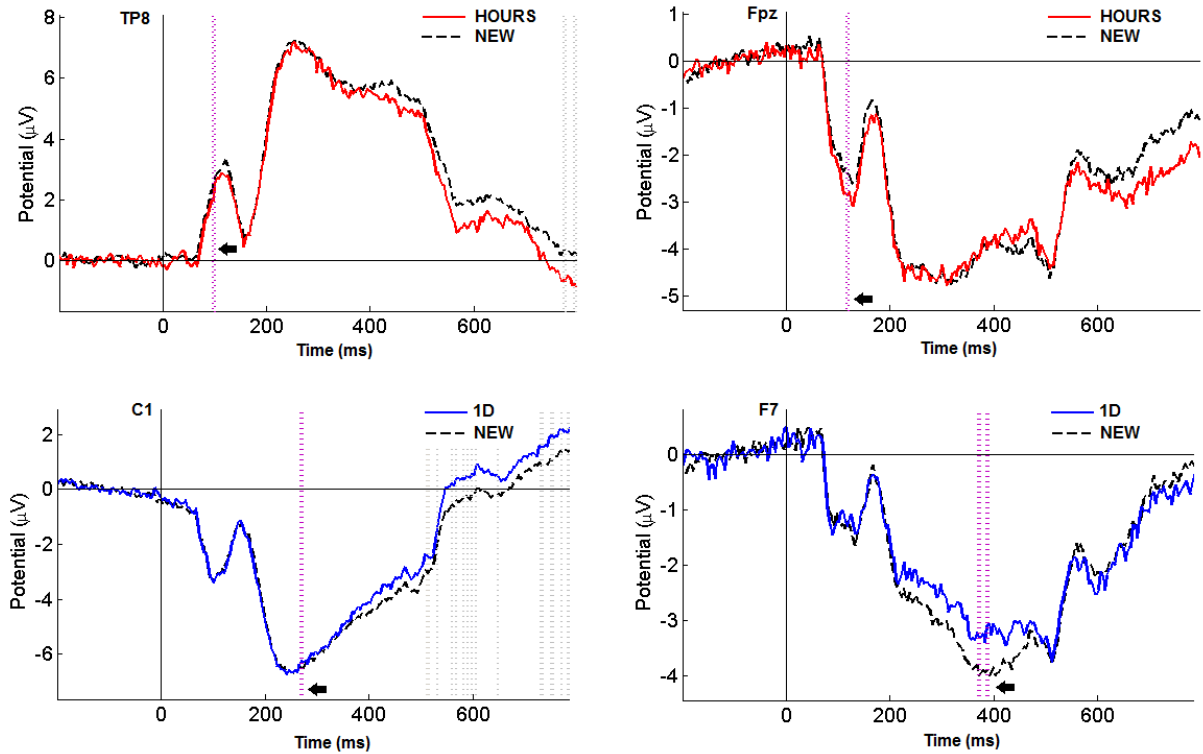


Figure 6.4 Early ERPs analysis on average across 11 participants for snapshots seen *HOURS* or *1D* before and compared to *New* pictures.

Four channels are presented: TP8 and Fpz for the contrast *HOURS* versus *NEW* and C1 and F7 for the contrast *1D* versus *NEW*. Significant *t*-test ($p < 0.05$, FDR corrected) for the time before stimulus offset ($t = -200$ -400 ms) are shown by vertical purple dashed lines and highlighted by a black arrow. Significant *t*-test ($p < 0.05$, FDR corrected) outside the time of interest are shown by vertical dashed grey lines.

3.2.3 Multivariate Pattern Analysis (MVPA)

We performed MVPA analyses on our EEG data by comparing the activity for snapshots seen in one of the 3 learning sessions (*HOURS*, *1D* or *3W*) and *NEW* snapshots. For each participant, classifier performance was measured for the trial tests at the same time point as for the training sets (diagonal matrix) or for any other time points (off-diagonal matrix). Performance was averaged across the 11 participants and calculated using the Area Under the Curve (AUC) where chance level corresponded to the same classification performed on random labels (Figure 6.5). The resulting matrix decoding for the 3 contrasts were squared-shape as found for *REMEMBER* versus *NEW* snapshots (Supplementary information Figure 6.S2) but that faded away over the retention intervals. Classification performance was found

significantly different from chance level from around $t = 500$ ms post-stimulus onset and until the end of the epoch for the comparisons HOURS/NEW, 1D/NEW and 3W/NEW.

Similarity analyses on the differential correct response percentages and on the decoding matrices were performed for each participant. Interestingly we found that the differential correct response percentages for the conditions HOURS/NEW, 1D/NEW and 3W/NEW was correlated with the respective decoding matrix performance (mean: $r = 0.57$, $SD = 0.7$, range: $-0.99 - 1.00$, Pearson's correlation coefficient). This suggests that it might be possible to infer participants' recall performance on very long term memories from the matrix decoding obtained through MVPA.

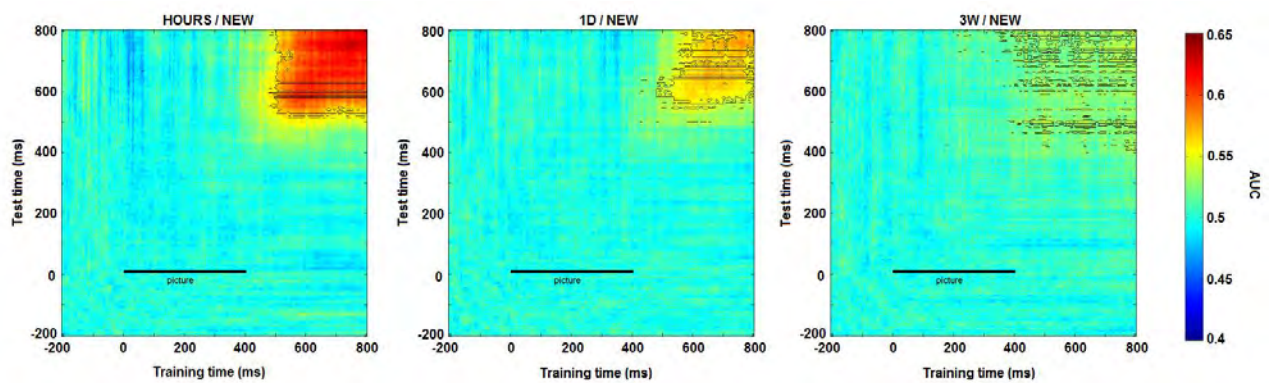


Figure 6.5 Decoding analysis for the comparisons between the 3 learning session (Hours, 1D, 3W) and New pictures on average across participants.

Significant values that lasted for a minimal duration of 30 ms are contoured in black (paired t -test, $p < 0.05$, FDR corrected).

4. Discussion

In this study we showed that MVPA techniques are sensitive enough to decode long-term memory traces. Whereas classical ERPs and Time-Frequency analyses were only able to differentiate memory traces of videos seen a few hours or one day before, MVPA allowed us to reveal a difference of pattern for snapshots seen 3 weeks before that was correlated with participants' performance. The dynamic of activation revealed by MVPA showed a sustained activity for the 3 retention intervals suggesting that a unique network of brain areas is involved in the retrieval of audiovisual content (J-R King and Dehaene, 2014). Whereas this activation is strong for stimuli recently encoded and corresponds to broad activations in centro-parietal region, the decrease of power for stimuli seen 3 weeks ago could be explained

by the involvement of a smaller brain network in frontal regions according to the ERP topography (not significant though) and/or in the hippocampus (Harand et al., 2012).

Theta band activations reflect long-range interactions that are linked to recollection processes (Doppelmayr et al., 1998; Klimesch et al., 1994; Nyhus and Curran, 2010; Osipova et al., 2006) and it has been found that an increase in the slow theta power (~3 Hz) correlates with recollection success (Pastötter and Bäuml, 2014). Intracranial recordings in humans have shown that slow theta oscillations are generated in the hippocampus to support the encoding of contextual details (Lega et al., 2016). These oscillations resemble the theta frequency band found in rodents between 4 and 10 Hz that play a critical role in spatial navigation (Jacobs, 2014). However, because in this study EEG recording was collected on the scalp of the participants, it is hard to make sure that this slow theta activity was generated by the hippocampus.

Although speculative, the decrease of power in the theta band over the retention intervals might reflect the involvement of a more localized pattern of neurons. However no significant activity was found in the gamma band during the retrieval of snapshots seen 3 weeks before, which is supposed to reflect more local activity (Nyhus and Curran, 2010; von Stein and Sarnthein, 2000). Sustained activity was strong for recent memories but weaker for longer retention intervals, the latter being associated with longer participants' reaction times. These results are consistent with the idea that memory recall can be interpreted in terms of diffusion model (Ratcliff and Mckoon, 2008; Ratcliff, 1978) whereby a decision is made based on the evidence accumulated over time and until a certain threshold. This perspective is already commonly held in the perception domain (Dehaene, 2009; Gold and Shadlen, 2007; Norris and Kinoshita, 2008; Ploran et al., 2007) and comes within the scope of the Bayesian brain (Knill and Pouget, 2004) but has not been so much investigated in the memory field (Koriat and Levy-Sadot, 2001; Moulin and Souchay, 2015; Shadlen and Shohamy, 2016; Wixted, 2009). In our case, we suggest that this decision criterion will be reached fast for recent memories whereas for older memories with a slight activation level, the process will take more time to be achieved. In agreement with this hypothesis we found that participants were faster to remember snapshots seen recently compared to the ones seen 3 weeks before. This suggests that remember responses emerge when a supposed-high criterion is reached. If activity is strong, but not sufficient to reach this criterion, judgment would correspond to Know response. In our experiment, failure to reach this criterion was associated with slow reaction times for Know responses. Although the differentiation between familiarity and recollection has been extensively reported and is supported by experimental evidence, it is not

clear yet whether they can be considered as variations of the same process (Dunn, 2004; Wixted, 2009) or emerge from distinct ones (Jacoby, 1991; Mandler, 1980; Yonelinas et al., 2010). Our data suggest a continuum between familiarity and recollection processes. It is important to point out that both participants' percentage of correct responses and reaction times for new responses were found in between the performance obtained for the condition 1D and 3W. This means that it might be easier for participants to compare new stimuli from old ones but that participants are better to detect recently seen stimuli than new ones. Participants' performance 3 weeks after the encoding phase was high which demonstrates that audiovisual clips are good material to study the recollection of long lasting memories though under-used in the literature (Ferguson et al., 2016; Furman et al., 2007; Tang et al., 2016). Importantly, sleep might have played a crucial role in the change of sustained activity found for the different retention interval. Indeed, numerous studies have shown that during sleep, a memory trace can be reactivated several times and therefore be consolidated into a stable neuronal pattern (Buzsáki, 1989; Girardeau et al., 2009; Peyrache et al., 2009; Sirota et al., 2003). However, in this design we could not isolate this variable to measure independently its impact.

Overall, we showed that long-term memory traces can be decoded based on single EEG trials. In this study 3 weeks was the longest retention interval. Knowing that memories can be maintained for several decades, it would be interesting to test whether decoding still applies for remote memories.

5. Conclusion

From this first EEG experiment I was happy to see that it was possible to differentiate pictures of audiovisual clips presented three weeks ago from pictures belonging to new clips. As I expected it, ERPs and time-frequency analyses were not powerful enough to detect slight differences for long retention intervals. Decoding techniques performed on single EEG trials were therefore necessary.

However, I am still surprised about the timing of these differences. Being in a lab that extensively studied recognition for very short latencies, it is quite disappointing to see that for a three week interval I was not able to see any early effect. Using decoding analyses I was only able to detect differences from about 400 ms after picture onset. For shorter retention intervals I found a couple of significant effects before 400 ms after picture onset. It is interesting to note that the earlier effects were found for images belonging to audiovisual clips seen just a few hours before and that for audiovisual clips seen one day before, early effects

occurred about 100 ms later. Such a delay might explain the differences of reaction time to recollect the memory of the audiovisual clips.

One of the benefits of decoding analyses is that we can get geometrical patterns that provide information concerning the dynamics of the underlying processing that we caught. I was quiet surprised to see that the same long lasting square pattern held for the three retention intervals. The difference was more in the power of the signal rather than in its nature. This suggested that the same neuronal pattern was at the basis of the retrieval process whatever the delay, at least until 800 ms after picture onset. I did not go even further, to make sure that the signal studied would not be “polluted” by any motor component. Given our experimental design and according to the subjective reports, it is very likely that this square pattern was at the basis of the recollection of the audiovisual clips.

6. Acknowledgements

This research was supported by the European Union’s Seventh Framework Programme (FP7/2007-2013)/ ERC grant agreement n°323711 to S.J.T. We thank Chris Moulin for helpful discussion and Estelle Bonin for her assistance in the stimuli preparation. The authors declare that there is no conflict of interest.

7. Supplementary information

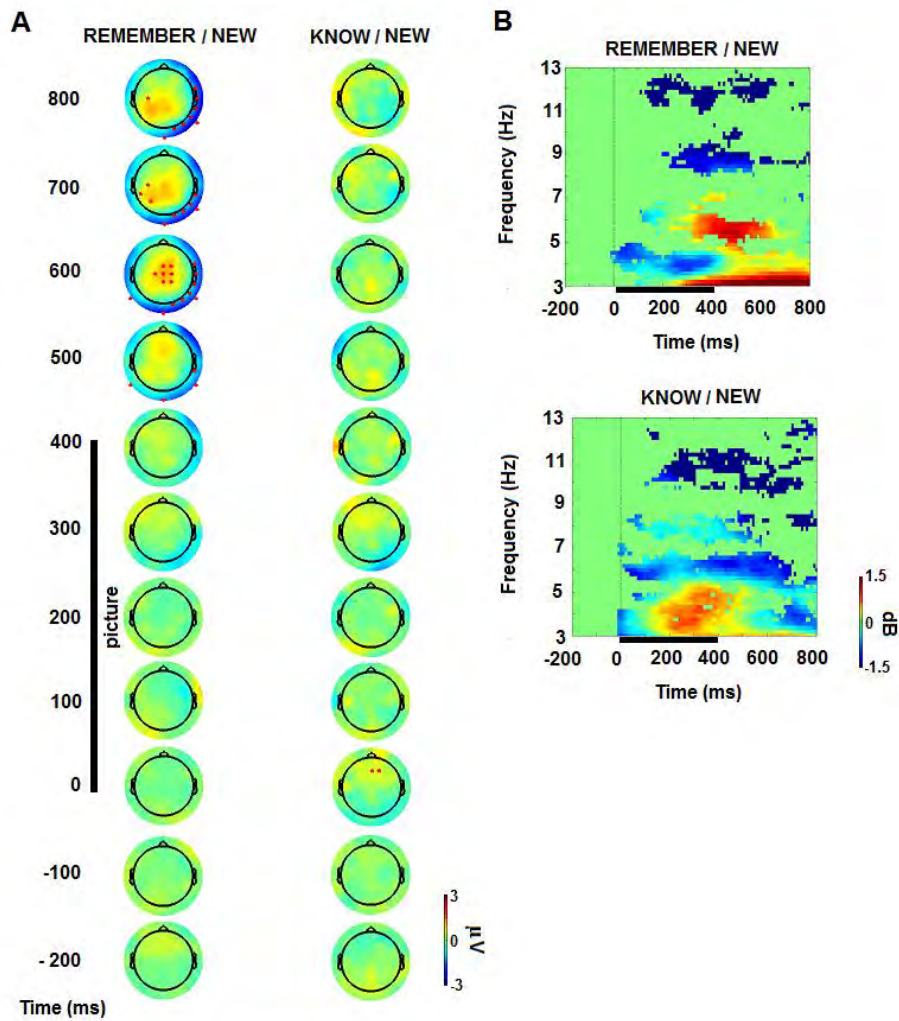


Figure 6.S1 Differential EEG activity between REMEMBER or KNOW snapshots and NEW snapshots.

Figures represent the average across the 11 participants for $t = -200-800$ ms post-stimulus onset. **A.** ERPs topography. Red dots correspond to significant channels ($p < 0.05$, FDR corrected). **B.** Time-frequency analysis. Only significant values (bootstrap $p < 0.05$, mask applied for 20 channels, all participants) are colored; non-significant values are plotted in green.

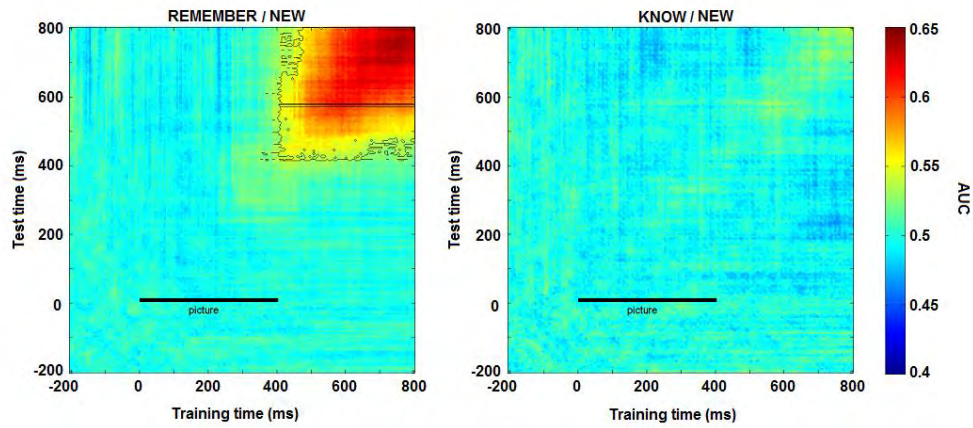


Figure 6.S2 Decoding analysis for the comparisons between Remember or Know responses and New pictures on average across participants.

Significant values that lasted for a minimal duration of 30 ms are contoured in black (paired t -test, $p < 0.05$, FDR corrected)

CHAPTER 7. Effect of repetition on explicit long-term memory: an EEG study

Chapter presentation

In the previous chapter the critical variable that I manipulated was the retention interval. However and as I developed it all along the manuscript, repetition might play an important role for retrieval in long-term memory. For short-term memory, the effect of repetition has been particularly studied with EEG but I could not find any article that tested its effect for longer delays such as a few weeks. I had this simple question in mind: what is the effect of repetition on the EEG signal during long-term retrieval?

The data from my first EEG experiment convinced me to that I had the tools to investigate this question. I would use the same experimental set up but this time, I would leave the retention interval fixed at 3 weeks and I would play around with the number of times the audiovisual clips would be presented. Simon and I agreed that the strongest effects would be seen for a couple of repetitions and that it was not necessary to play the same audiovisual clip ten times. So we decided to present the clips either once, twice, three or five times.

After having done my first EEG experiment, I knew that I would get data that I could exploit in about 6 months. I could use again most of the scripts that I wrote to run the experiment and the analysis of the behavioral and the EEG data. Everything was well set for me to hire an intern on this project. To me, this was an important part of my PhD. I was an intern myself just a couple of years before and I knew how much I gained from the PhD students I worked with. Simon was happy with this idea and gave me a free hand to hire anyone who would fit. One interview was enough for me: Meyi was excited by the project and full of enthusiasm. She will be the one to work with me on this topic.

1. Introduction

“Every one knows that repetition plays an important part in the process of acquiring knowledge, but hitherto there has been no attempt experimentally to study this factor beyond the experiments of Ebbinghaus relating to the effect of repetition on the duration of memory” (Smith, 1896).

This sentence was written by W.G. Smith at the end of the nineteenth century but sounds very relevant today and especially when considering explicit long-term memory. Explicit memory reflects an intentional attempt to recollect an event. It is traditionally viewed as being the result of two potential processes: a recollection process when a memory is retrieved with contextual details or based on a familiarity judgment (Mandler, 2008, 1980; Yonelinas, 2002). These two subjective judgments can be assessed using a Remember/Know/New paradigm in which remember and know responses reflect respectively recollection and familiarity even if this equivalence might not be always true (Gardiner and Richardson-Klavehn, 2000; Williams et al., 2013). Two different Evoked Response Potentials (ERPs) components have been identified to characterize these processes: the FN400, a midfrontal negativity occurring within 300-500 ms after stimulus onset which corresponds to familiarity judgments and the LPC, a late positive component, that kicks in between 500 and 800 ms in the left parietal region and that reflects recollection processes ((Curran, 2000; Duarte et al., 2004; Düzel et al., 1997; Tsivilis et al., 2001) and for a review see (Rugg and Curran, 2007)). In addition to differences of ERP latencies, an increase of power in the theta band (4-7 Hz) and in the gamma range (30-100 Hz) characterizes respectively recollection and familiarity judgments (Gruber and Tsivilis, 2008). Evidence showed that both remember and know responses can be influenced by the number of presentations an item was presented during learning. In particular, the repeated presentation of an item is able to enhance remember responses (Dewhurst and Anderson, 1999; Gardiner et al., 1996; Mäntylä and Cornoldi, 2002). This repetition effect corresponds to an increase of activity in the parietal region (de Chastelaine et al., 2009; Friedman et al., 2010; Nessler et al., 2007) that was associated with shorter latencies: from 300 to 500 ms (de Chastelaine et al., 2009; Friedman et al., 2010) and therefore called “early parietal old/new effect” by de Chastelaine et al. Enhanced theta activity was also found for a repeated item compared to a new one (Klimesch et al., 2000) but with no difference across the number of presentations (Van Strien et al., 2005). Whereas it has been shown that repetition might affect explicit memory for several decades (Bahrick et al., 1975) EEG studies mainly

focused on the effect of repetition for very short retention intervals between the learning and the explicit retrieval of the repeated stimuli. To our knowledge, no EEG data exist for retention intervals exceeding a day. This is important to point out because the consolidation of a memory trace which involves both a synaptic and a system-level consolidation might take several weeks to be completed (Frankland and Bontempi, 2005). This gradual stabilization can be even picked out in some situations with the increase of performance after a day of sleep (Walker and Stickgold, 2010). Repetition effect might be therefore different for longer retention times.

Here we decided to test the effect of repetition during the recall of short audiovisual clips three weeks (experiment 1) and five months (experiment 2) after the learning phase while participants' EEG activity was recorded. Note that the data have not been yet fully analyzed but already provide interesting information.

EXPERIMENT 1

2. Material and Methods

2.1 Participants

Twelve healthy subjects (mean age = 24.5 y, SD = 2.9, 6 females) gave written informed consent and participated to the experiment approved by the INSERM Ethical Evaluation Committee (CEEI).

2.2 Stimuli

Overall, 840 4-seconds audiovisual clips (size = 320*240) and downloaded from the internet were used in the experiment. Among them 120 clips of Michael Jackson were only used for controlling participants' attention during learning sessions and were not used in the recall test. The remaining 720 audiovisual clips were selected to be as diversified as possible and unknown from the participants. Two-third of them (n = 480) were shown during learning sessions and were counterbalanced between participants. One frame of each of the 720 clips was selected to be presented during the recall phase as a video snapshot. The same 720 snapshots were presented for all the participants.

2.3 Task

The experiment included two phases: a learning phase and a test 3 weeks later to assess long-term recall (Figure 7.1).

During the learning phase, 480 audiovisual clips were shown to the participants on a computer screen. The audiovisual clips were presented either once, twice, three or five times with the same number of clips by presentation condition ($n = 120$). Because of the massive number of trials, the learning phase was split into two independent sessions performed on two consecutive days where half of the audiovisual clips were presented. Participants were instructed to pay as much attention as possible to all the video clips and were informed that they will be tested on the recall of these videos three weeks later. To control for participants attention during viewing, subjects were asked to make a button-press response whenever Michael Jackson appeared on an audiovisual clip ($n = 60$ by session). Participants could also indicate that they already knew a clip by pressing a key assigned to this purpose during the clip presentation and the corresponding trials were discarded from further analysis. The duration of a learning session was about 1.30 h.

The recall test performed three weeks after the first learning session (mean = 21days, 2h37min, SD = 14h02 min) was a cued-recall task of 720 trials and during which EEG's participants was recorded. Memory recall was assessed with a classical Remember/Know/New judgment (Tulving, 1985a). Each trial started with a fixation cross of varying duration (600-1000 ms) followed by either a snapshot of a clip seen during the learning sessions ($n = 480$) or a new one (NEW) ($n = 240$) that was presented for 400 ms. After picture offset, participants were asked to press as fast and as accurate as possible, with the same hand, one of the three keys corresponding to Remember, Know and New judgments (randomized between participants). Before testing, participants were informed of the meaning of these three response possibilities as follows: 'Remember': you identify the picture as previously seen in one of the two learning sessions and you can replay the related-video in your mind by anticipating the following scenes of the short scenario. 'Know': you identify the picture as previously seen in one of the two learning sessions but you can't retrieve to which video it belongs to. 'New': the picture does not belong to any of the videos shown during the two learning sessions.

Participants were then invited to judge their response on a five-point confidence scale (1: "Not sure at all"; 2: "A bit sure"; 3: "Fairly sure"; 4: "Very sure"; 5: "Completely sure"). The

duration of the recall phase was about 1h15. Learning and recall phases were programmed using the Psychtoolbox, a Matlab toolbox.

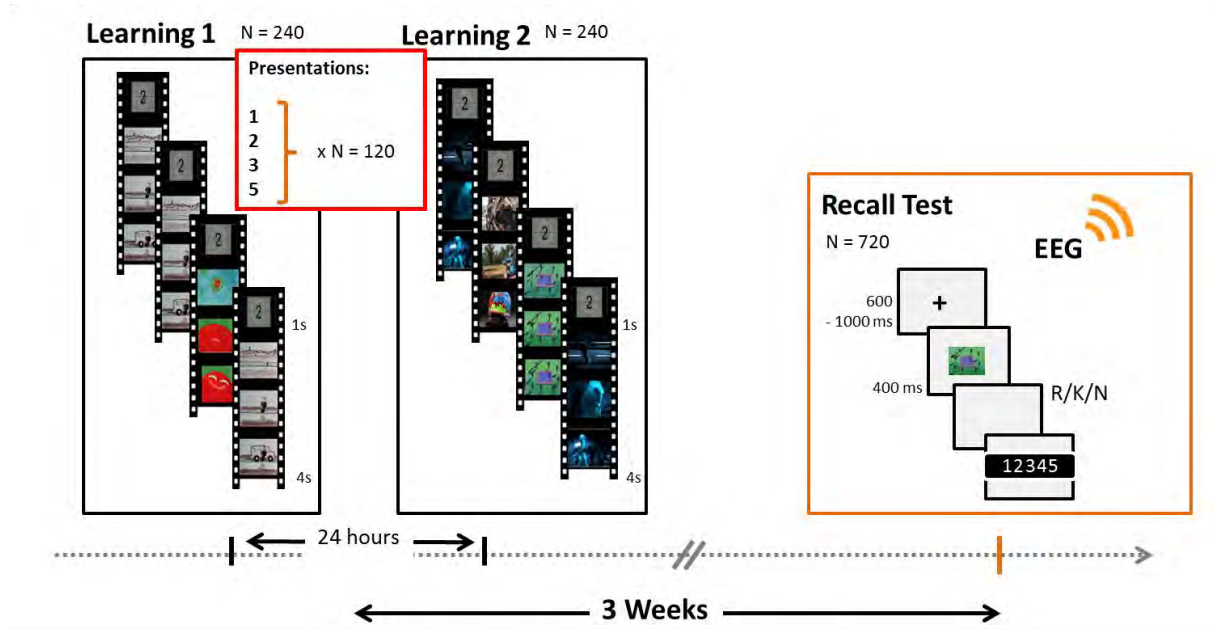


Figure 7.1 Experimental design.

Every participant watched 480 different audiovisual clips during 2 learning sessions with a 24-hour time interval between the two. A recall test was performed 3 weeks after learning. For each learning session 240 short audiovisual clips (3s) were presented either once, twice, three or five times on screen. Over the 2 learning sessions, participants watched 120 different clips for a given number of presentations. Memory test was a cued-recall in which snapshots of clips seen during the learning sessions ($n = 480$) or new ones ($n = 240$) were briefly presented on screen (400 ms). After picture offset, participants were invited to make as fast and as accurate as possible a R/K/N judgement on the picture and rate the confidence on their response on a 5-point scale. EEG was only recorded during the recall phase.

2.4 EEG acquisition and analysis

All the 12 participants were included in the EEG analysis.

2.4.1 EEG acquisition and preprocessing

During the recall phase, participants' brain activity was recorded with a 64-channel cap connected to a BioSemi EEG amplifier (5 kHz) and with a sampling frequency of 1024 Hz. Data were filtered with a finite impulse response Blackman band-pass filter with a lower cutoff frequency of 0.1 Hz to avoid filtering artifacts (Acunzo et al., 2012) and a higher cutoff

frequency of 100 Hz with a transition bandwidth = 1 and then down sampled to 256 Hz. Trials were then epoched from -1s to 2 s relatively to picture onset, their baseline [-0.2 s 0] removed and re-referenced based to average reference. Artefacts were removed by performing ICA. For all participants, trials with extreme values for a channel ($>100 \mu\text{V}$ or $< -100 \mu\text{V}$) or abnormal trend (slope max: $75 \mu\text{V}$, $R^2 > 0.3$) or abnormal distribution (kurtosis outside 5 standard deviations of the kurtosis for each single electrode) were excluded from the analysis after visual inspection. All the preprocessing steps were done with EEGLab (Delorme and Makeig, 2004).

2.4.2 ERPs analysis

Trials were split according to the five numbers of presentations: ONE, TWO, THREE, FIVE and NEW for snapshots respectively seen once, twice, three times and five times or never seen in the learning sessions, or according to participants subjective responses: REMEMBER, KNOW and NEW responses. Differential activity was measured on a two by two analysis by contrasting one condition with a NEW condition (retention interval or subjective response) and statistical difference was performed using Student t-test with a statistical threshold of $p < 0.05$ after False Discovery Rate (FDR) to correct for multiple comparisons.

2.4.3 Decoding

Multivariate pattern analyses were conducted on the same data used for the ERPs analysis. The classification was performed on a two by two analysis with one condition constituted by trials of NEW snapshots and the second condition containing snapshots seen for a specific number of times (ONE, TWO, THREE or FIVE). Trials were randomly split 120 times according to the Monte-Carlo cross-validation procedure: 120 Cross Validations (CV). For each cross-validation (CV), the number of trials was equalized between conditions 90% of the trials were used by the algorithm as a training set, the remaining 10% as a test set. Signal was normalized for the training and test sets. A linear classifier (L2 –regularized logistic regression, (Fan et al., 2008)) was trained on individual trials at each time point from -200 to 800 ms and tested on individual trials on the same and all the other time points (temporal generalization). The performance of the classifier for every participant and at each test time point was evaluated by using the Area Under the Curve (AUC). The average decoding performance across participants resulted in a matrix decoding where the x- and y-axes respectively represented training and testing times. Chance level decoding was calculated by performing the same classification on randomly permuted labels. Statistical significance was

reached when the average across the 12 participants was higher than chance level for a duration of at least 30 ms of the training or testing time (paired t-test, $p < 0.05$, FDR corrected). Decoding analyses were performed by using an open source script developed by Sebastien Crouzet that can be downloaded at <https://github.com/scrouzet/classifyEEG> and that was used in (Crouzet et al., 2015) and in (Cauchoix et al., 2016).

3. Results

On average, five out of the 480 videos ($SD = 5.0$) were considered already known by participants and the corresponding trials were discarded from behavioral and EEG analysis for each subject.

3.1 Behavioral results

3.1.1 Learning phase

Participants watched the audiovisual clips during two learning sessions of respectively 240 clips each that were done on two consecutive days. To make sure that participants were attending the videos, they were asked to press the space bar whenever a video of Michael Jackson was presented (60 videos by learning session). The high percentage of HITS and the low False Alarm (FA) rate for the two learning sessions show that participants were paying attention to the videos presented on screen (session 1: HITS = 96.4% ($SD = 6.8$); FAs = 0.01% ($SD = 0.04$) and session 2: HITS = 98.7% ($SD = 2.0$); FAs = 0.09 % ($SD = 0.2$).

3.1.2 Recall phase

During the recall phase performed three weeks after the first learning session, snapshots from the clips seen by the participants: the TEST pictures ($n = 480$) or NEW pictures ($n = 240$) were briefly presented on screen. Participants were asked to make a Remember-Know-New judgment as fast and as accurate as possible and to rate the confidence on their response on a five-point scale.

- Correct response analyses

On average participants were 59.8% ($SD = 6.4$, range: 46.9% - 69.9%) correct at discriminating pictures as new or already seen with ‘Remember’ and ‘Know’ responses pooled together. This shows that on average participants were able to recall short audiovisual clips seen three weeks ago based only on the presentation of single snapshots. This reflects in the average d' score as well ($d' = 1.0$, $SD = 0.15$). HITS corresponded to ‘Remember’ or

‘Know’ responses for snapshots previously seen during the learning sessions. CRs corresponded to snapshots correctly identified as never seen in the learning sessions: ‘New’ response. The percentage of HITs increased over the number of presentations: 36.4% (SD = 14.7), 48.6% (SD = 17.4), 58.4% (SD = 17.3) and 69.6% (SD = 16.2) for 1, 2, 3 and 5 presentations respectively and the percentage of CRs was high: 72.7% (SD = 15.7) and was even higher than the percentage of correct responses for pictures seen once or twice (one-way ANOVA: $F(4,55) = 10.2$, CI_{95} -Difference (‘1’ vs ‘2’) = [-31.0 to 6.6], CI_{95} -Difference (‘1’ vs ‘3’) = [-40.8 to -3.2], CI_{95} -Difference (‘1’ vs ‘5’) = [-52.0 to -14.4]), CI_{95} -Difference(‘New’ vs ‘1’ = [17.6 – 55.1], CI_{95} -Difference(‘New’ vs ‘2’ = [5.4 – 42.9]) (Figure 7.2A). However, we did not find any difference in participants’ reaction times for all of these conditions. The median reaction times for HITs over the number of presentations was: 1.39 s (SD = 0.33), 1.36 s (SD = 0.29), 1.34 s (SD = 0.26) and 1.24 s (SD = 0.19) for snapshots respectively seen from 1 to 5 times and for CRs concerning the new pictures: 1.19 s (SD = 0.23) (Figure 7.2B). The increase of the percentage of recall over the number of presentations was combined with the confidence increase from ‘Medium sure’ = 3 to ‘Very sure’ = 4 as shown by the average confidence score from 1 to 5 presentations respectively: 3.0 (SD = 0.5), 3.5 (SD = 0.5), 3.6 (SD = 0.6) and 4.0 (SD = 0.5) calculated on correct responses only (one-way ANOVA: $F(3,44) = 6.68$, CI_{95} -Difference (‘1’ vs ‘2’) = [-1.0 to 0.1], CI_{95} -Difference (‘1’ vs ‘3’) = [-1.1 to -0.0], CI_{95} -Difference (‘1’ vs ‘5’) = [-1.5 to -0.4], CI_{95} -Difference (‘2’ vs ‘3’) = [-0.7 to 0.4], CI_{95} -Difference (‘2’ vs ‘5’) = [-1.0 to 0.1], CI_{95} -Difference (‘3’ vs ‘5’) = [-0.9 to 0.2]). This shows that the number of presentations has an effect on metacognitive processes and here in particular, on participants’ ability to evaluate correctly their response.

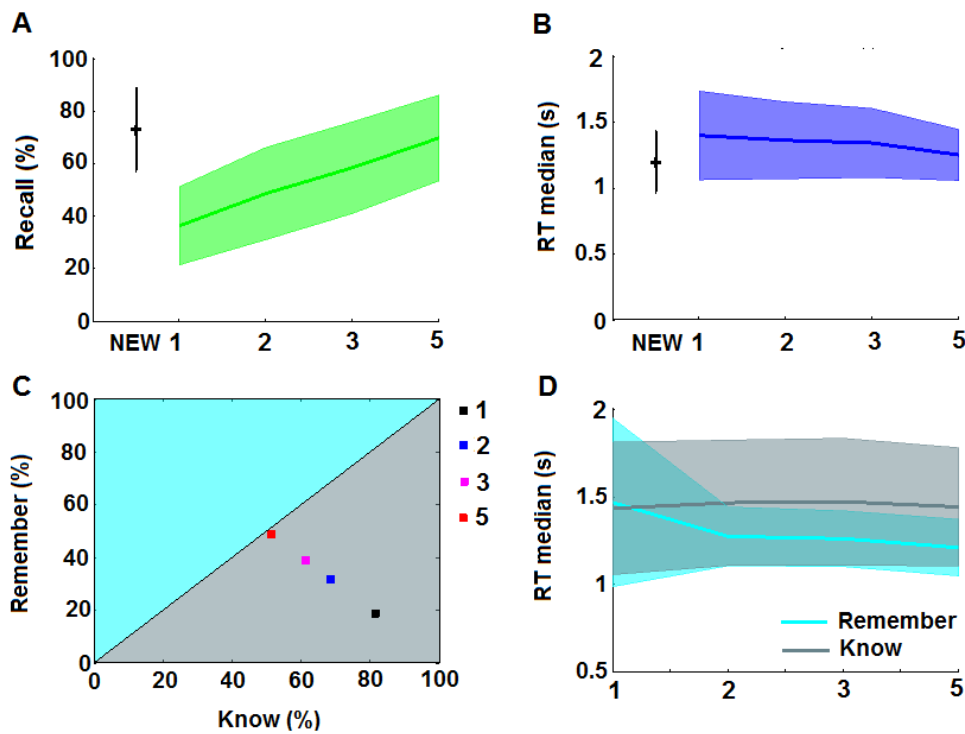


Figure 7.2 Participants recall performance for videos seen 1, 2, 3 or 5 times 3 weeks before. *A.* Recall percentage for snapshots never seen, seen once, twice, three or five times (mean and standard deviation). *B.* Reaction times (s) for correct responses over the number of presentations during learning (median and standard deviation). *C.* Repartition of Remember and Know responses (%) for HITs according to the number of exposures during learning. *D.* Reaction times (s) for correct Remember and correct Know responses over the four number of presentations (median and standard deviation).

- Correct Remember and Know analysis

For the pictures that were correctly identified as previously seen, 62.8% (SD = 20.2, range: 32.9 – 96.4) of them corresponded to ‘Know’ responses, which was higher than the number of ‘Remember’ responses ($t(22) = -3.10$, $CI_{95}\text{-Difference}$ (“Remember” vs “Know”) = [-42.8 to -8.5]). Interestingly, the ratio of correct ‘Remember’/ correct ‘Know’ increased over the number of presentations: 18.8% (SD = 16.8), 31.7% (SD = 19.0), 38.8% (SD = 22.21), 48.6% (SD = 25.6) for snapshots respectively seen once, twice, three times and five times (one-way ANOVA: $F(3,44) = 4.24$, $CI_{95}\text{-Difference}$ (“ratio for ‘1’ ” vs “ratio for ‘2’ ”) = [-36.0 to 10.1] $CI_{95}\text{-Difference}$ (“ratio for ‘1’ ” vs “ratio for ‘3’ ”) = [-43.1 to 3.0], $CI_{95}\text{-Difference}$ (“ratio for ‘1’ ” vs “ratio for ‘5’ ”) = [-52.9 to -6.8], $CI_{95}\text{-Difference}$ (“ratio for ‘2’ ” vs “ratio for ‘3’ ”) =

[-30.2 to 15.9], CI_{95} -Difference (“ratio for ‘2’ ” vs “ratio for ‘5’ ”) = [-40.0 to 6.1], CI_{95} -Difference (“ratio for ‘3’ ” vs “ratio for ‘5’ ”) = [-33.3 to 13.3]).

This demonstrates a long term effect of the number of presentations which directly affects participants’ subjective experience when they try to recall the videos (Figure 7.2C and Figure 7.S1).

Interestingly, reaction times decreased for correct Remember responses. In contrast, no difference was found for correct ‘Know’ responses over the number of presentations

The median reaction times for correct Remember were: 1.46 s (SD = 0.5), 1.27 s (SD = 0.2), 1.26 s (SD = 0.2) and 1.21 s (SD = 0.2) for snapshots seen respectively once, twice, three or five times (figure 7.2D). The median reaction times for correct Know responses for the same number of exposure were: 1.43 s (SD = 0.4), 1.46 s (SD = 0.35), 1.47 s (SD = 0.4), 1.44 s (SD = 0.4). No main difference of reaction time was found for correct Remember or correct Know responses depending on the number of exposures however we found that Remember responses were faster than Know responses (two-way ANOVA: Recall: $F(1,87) = 5.13$, $p < 0.05$; Repetition: $F(3,87) = 0.63$; Recall * Repetition: $F(3,87) = 0.88$). ‘Remember’ responses were faster than ‘Know’ responses but this does not mean that familiarity was processed after recollection. It is due to our experimental design that pushes participants to make preferentially fast ‘Remember’ responses. Late ‘Know’ responses simply indicate that they were done when recollection failed after some time, about 1.5 s on average.

Overall, our behavioral results indicate that participants can explicitly retrieve the memory of videos seen three weeks before based on the presentation of single snapshots. Importantly, the number of exposures to the stimuli affects directly recall processes as shown by differences in terms of percentage of recall and confidence in the response. What is the underlying neuronal mechanism that drives such differences? In the next section we will analyze participants’ EEG signal during recall. We will present ERPs and decoding analyses.

3.2 EEG analysis

3.2.1 Differential ERP analysis

Trials corresponding to the four numbers of exposures: ONE, TWO, THREE and FIVE were compared with trials of NEW pictures. ERPs analysis was performed on the mean average across participants for $t = -200$ -800 ms after picture onset. Significant channels were found for a threshold at $p < 0.05$ (paired t-test, FDR corrected). The ERP topography (Figure 7.3)

shows a significant difference for the condition FIVE/NEW for $t = 600-700$ ms after picture onset mostly in the central and parietal electrodes. This topography was representative of the contrast Remember/NEW performed independently of the number of times the audiovisual clips were seen (Supplementary information, Figure 7.S1A). This component might be a marker of the recollection of audiovisual clips. For the contrast ONE/NEW, TWO/NEW, THREE/NEW and FIVE/NEW no significant channels were found within the whole epoch.

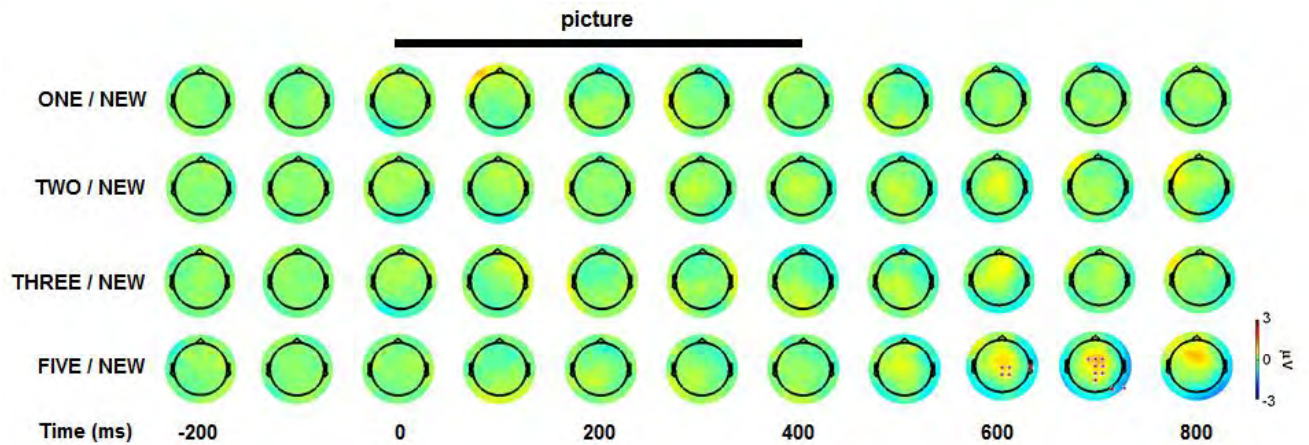


Figure 7.3 Differential EEG activity between snapshots seen ONE, TWO, THREE or FIVE TIMES, three weeks ago and NEW snapshots.

The ERP topography represents the average across the 12 participants for $t = -200-800$ ms post-stimulus onset. Red dots correspond to significant channels ($p < 0.05$, FDR corrected).

3.2.2 Multivariate Pattern analysis (MVPA)

We performed MVPA analyses on our EEG data by comparing the activity for snapshots seen ONE, TWO, THREE or FIVE times three weeks ago and NEW snapshots. For each participant, the classifier performance was measured for the trial tests at the same time point as for the training sets (diagonal matrix) or for any other time points (off-diagonal matrix). Performance was averaged across the 12 participants and calculated using the Area Under the Curve (AUC) where chance level corresponded to the same classification performed on random labels (Figure 7.4).

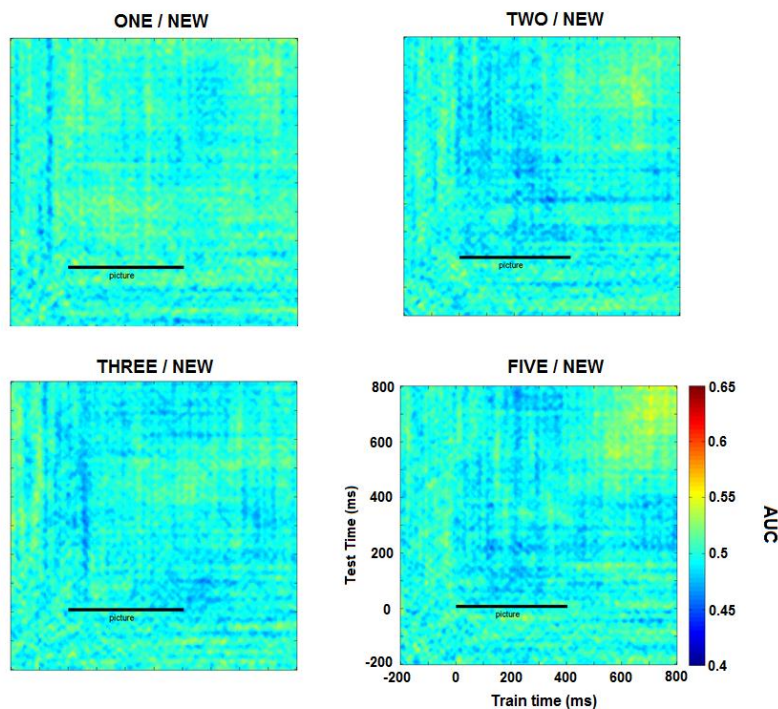


Figure 7.4 Decoding analyses for the comparisons between the four numbers of exposures (ONE, TWO, THREE, FIVE) and NEW pictures on average across participants.

Note that no statistical test was performed

The resulting matrix decoding for the audiovisual clips seen five times compared to new clips were squared-shape as found for the data REMEMBER versus NEW in chapter 6 (Supplementary information Figure 7.S2). It is important to note that for five participants, the decoding matrix for the condition FIVE/NEW was more pronounced than the effect found over the 12 participants (Supplementary information Figure 7.S2). It would be interesting to see whether decoding performance correlate with participants' performance, as found in the previous chapter.

4. Discussion

Overall, participants' performance to recall the audiovisual clips was lower than what we could expect from chapter 6. In the previous chapter, participants were able to identify 64.5% (SD = 18.1) of the audiovisual clips that were seen twice and three weeks ago. In the current experiment only 48.6% (SD = 17.4) of the audiovisual clips seen twice were correctly identified. This difference could be explained by several parameters that varied across the experiments: the number of audiovisual clips that were shown, the duration of each audiovisual clip and the number of items presented by learning session. Indeed, in the current

experiment, we presented 480 clips of a 3-s duration in two sessions whereas in the previous chapter, 375 clips of a 7-s duration were presented within three learning sessions. In addition, we did not control for the spacing effect, the number of video clips between two presentations of the same audiovisual clip, an effect that influences long-term retrieval (Cepeda et al., 2006). While the ERP topography could only differentiate snapshots seen five times from new snapshots, MVPA analyses were not better. Surprisingly, the algorithm was not able to differentiate snapshots seen two or three times from new snapshots. Given the behavioral and the decoding performance, one main question arises: Is the memory trace of all of these audiovisual clips lost? Or are participants not able to re-access the memory trace? The cued-recall task might be not sufficient to elicit participants' retrieval. Indeed, it is difficult to retrieve the memory of audiovisual clips from a still frame. However, given the behavioral data from chapters 3 and 4, it would be very surprising to find that participants would not be able to recognize these audiovisual clips, even several months after. In the following experiment, we tested the same participants five months later, to see whether they could recognize the audiovisual clips from new ones. Note that the data have not been fully analyzed. For now, we will only provide the first behavioral results of this experiment.

EXPERIMENT 2

5. Material and Methods

5.1 Participants

Seven healthy subjects (mean age = 25y, SD = 3.5, 3 female) out of the 12 participants who did the experiment 1 accepted to participate in this second experiment that was approved by the INSERM Ethical Evaluation Committee (CEEI).

5.2 Stimuli

In addition to the 480 audiovisual clips used in experiment 1 and learnt in the two learning sessions: OLD clips, 480 new audiovisual clips were used in this experiment. These audiovisual clips were also 3-s long and downloaded from the internet. The first frame of each of the 960 clips was also selected to be presented during the recall phase. The same snapshots were presented for all the participants.

5.3 Task

The experiment consisted in a second testing phase performed after the experiment 1 and was done five months later after the first learning session (Figure 7.5).

Each trial ($n = 480$), started with a fixation cross that was associated after 600 to 1000 ms with the display of the first 800 ms of the audio of an OLD ($n = 240$) or NEW ($n = 240$) clip. After the offset of the sound, the fixation-cross remained on screen during 200ms. The first frame of the clip corresponding to the previous sound was then displayed for 400ms. Note that the picture presented on screen always matched the sound presented first, either for OLD or NEW clips. Participants were asked to press as fast and as accurately as possible with the same hand one of two keys to decide whether the picture comes from an OLD or a NEW clip (key were randomized across participants). After participants' answer, two videos were successively presented and corresponded to the first 2-s of an OLD and NEW clip. One of these clips was the one from which the sound track and the snapshot that were displayed at the beginning of the trial were taken from. After the presentation of the second video, participants were asked to press with the same hand one of two keys to report whether it was the first clip or the second clip that they had seen in the lab during the two learning session, five months ago. The recall phase lasted about 2h30 and was programmed on Psychtoolbox, a Matlab toolbox.

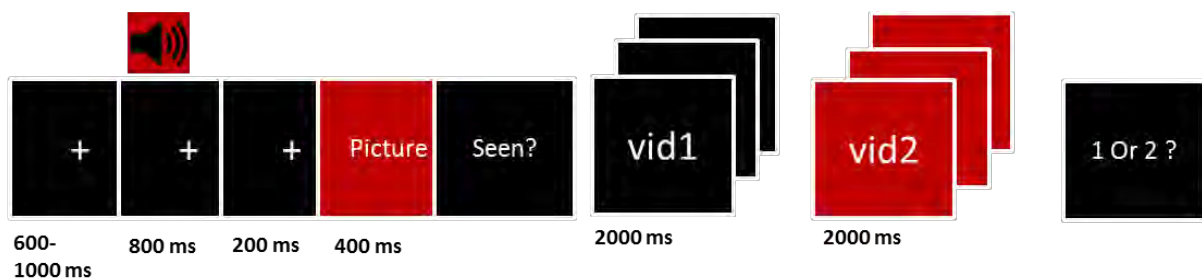


Figure 7.5 Experimental design.

For each trial ($n = 480$) the first 800 ms of the sound track of an audiovisual clip was displayed before the presentation of the first frame of the corresponding audiovisual clip for 400 ms. After picture offset, participants were invited to say as fast and as accurately as possible whether the picture corresponded to an OLD audiovisual clip ($n = 240$) or to a NEW audiovisual clip ($n = 240$). After participants' response, the first 2-s of two audiovisual clips were successively presented on screen. The audiovisual clip from which the sound track and the snapshot were displayed at the beginning of the trial was necessarily one of these two clips (in red in the figure) and could be either an OLD audiovisual clip ($n = 240$) or a NEW

clip ($n=240$). After the presentation of the second clip, participants were asked to identify which of the two videos was an OLD audiovisual clip. Participants' EEG activity was recorded during the task ($n = 7$).

6. Results

6.1 Behavioral results

Five months (mean = 5 months and 20 days, $SD = 7$ days) after the first learning session of the experiment 1, participants were first asked to say whether the picture presented on screen corresponded to an OLD or NEW picture. Participants were then invited to identify from two clips which one was an OLD clip.

Participants were on average 60.02% ($SD = 5.2$, range: 52.3 – 68.3%) correct in identifying a picture as either OLD or NEW whereas they were 73.7% ($SD = 8.1$, range: 62.7 – 87.7%) correct in identifying the OLD clip from a NEW one. For the OLD pictures, participants were 47.4% ($SD = 13.4$), 53.6% ($SD = 22.5$), 44.6 % ($SD = 21.1$) and 45.9% ($SD = 25.5$) correct for audiovisual clips seen respectively once, twice, three times and five times. Participants were 72.3% ($SD = 12.6$) correct in identifying NEW pictures. Concerning the audiovisual clips, participants were 70.6% ($SD = 7.6$), 78.8% ($SD = 8.7$), 69.2% ($SD = 11.8$), 76.3% ($SD = 13.0$) correct in identifying the OLD clips from NEW ones for the audiovisual clips seen respectively once, twice, three times and five times. A two-way ANOVA analysis showed that participants' performance was not different over the number of exposures but that participants were better to recognize the audiovisual clips presented as such than when only cued by their first frame (two-way ANALYSIS: Stimuli: $F(1,48) = 33.64$, Exposures: $F(3, 48) = 0.84$, Stimuli*Exposures: $F(3,48) = 0.13$)).

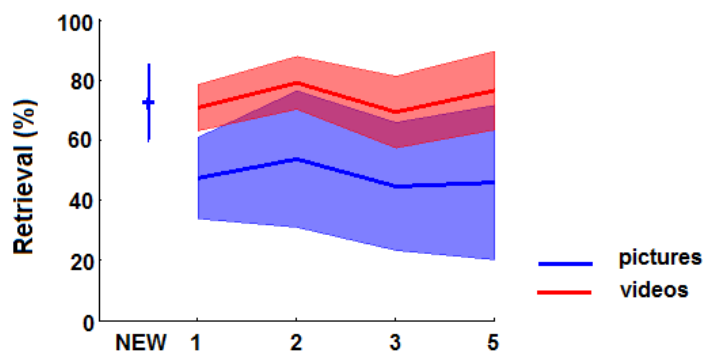


Figure 7.6 Retrieval performance over the number of exposures for audiovisual clips seen five months ago.

Retrieval was based on the presentation of the first frame of the audiovisual clip (in blue) or on the 2-s presentation of the clip (in red). The condition NEW was only measurable for retrieval based on the picture. Retrieval is expressed in % \pm SD.

7. Discussion

The preliminary results of the experiment 2 showed that the memory trace of the audiovisual clips presented five months earlier was still accessible to the participants. However we found that the presentation of a single still frame was not enough to elicit the retrieval of the audiovisual clips, even for the ones presented five times. A stronger cue was needed to bring the memory back to mind. Here we show that the presentation of the first 2 s of the 3-s audiovisual clips in a forced-choice task was enough for participants to be able to recognize an OLD clip from a NEW one. Although participants' performance for the OLD pictures was at chance level, the participants were good at categorizing the NEW pictures. Indeed participants were 72.3% (SD = 12.6) correct for the NEW pictures, a result that is similar to the percentage of correct responses to NEW pictures found in experiment 1: 72.7% (SD = 15.7) and in chapter 6: 79.7% (SD = 13.3). The ease to detect novel items might have helped participants to choose the correct audiovisual clip in a forced-choice task. It is therefore difficult to quantify to what extent the presentation of the audiovisual clips itself provided more information than the still frames. Further investigations should compare participants' performance in two forced-choices, one with still frames and one with audiovisual clips.

The reason we did not directly compare participants' performance in two forced-choice tasks is that we wanted to increasingly provide information about the audiovisual clips to elicit retrieval. This second experiment was designed to highlight the gradual accumulation of evidence until participants' decision. Accordingly, we displayed at the beginning of the trial the first 800 ms of the audiovisual clip. Although this cue seems too weak to elicit the retrieval of the whole audiovisual clip it might already trigger a search in memory. This search can be continued with the presentation of the still frame corresponding to the sound. As shown by participants' performance, the succession of these two cues was too weak to elicit the recall of the clips. However information about the clip might have been accumulated and could be potentially detectable on the EEG signal. If it is the case, we would be able to predict from this signal if participants can identify the correct clip in the forced-choice.

8. Supplementary information

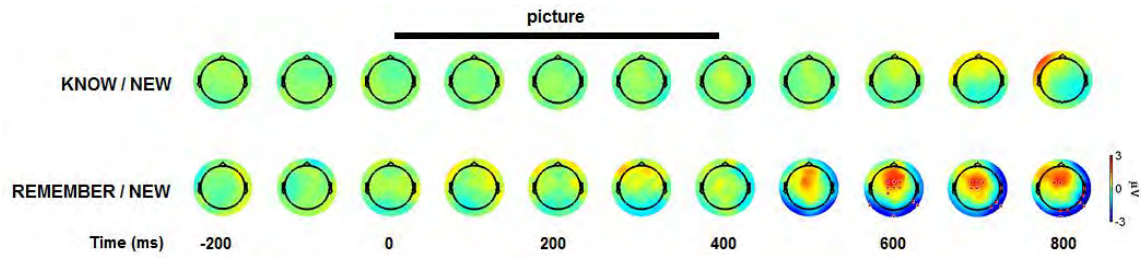


Figure 7.S1 Differential EEG activity between REMEMBER or KNOW snapshots and NEW snapshots.

The ERP topography represents the average across the 12 participants for $t = -200-800$ ms post-stimulus onset. Red dots correspond to significant channels ($p < 0.05$, FDR corrected).

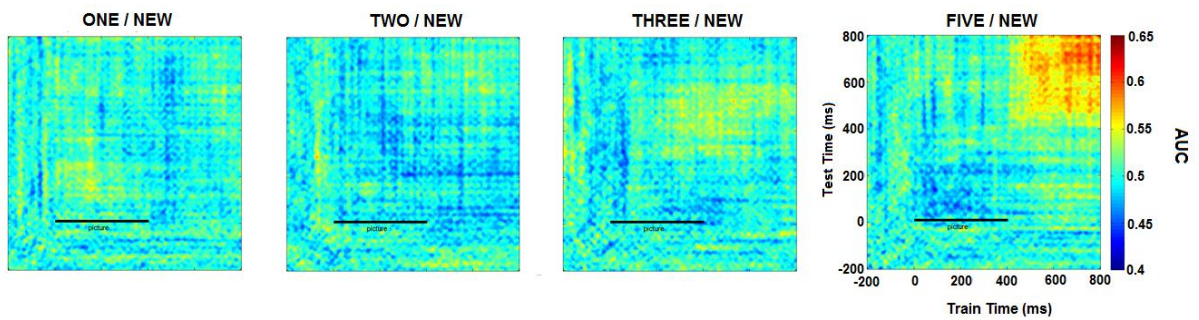


Figure 7.S2 Decoding analyses for the comparisons between the four numbers of exposures (ONE, TWO, THREE, FIVE) and NEW pictures on average across the five participants with the best decoding performance.

Note that no statistical test was performed

GENERAL DISCUSSION

1. Wrap-up

1.1 Overall framework

Since the 60's there was increasing evidence that consolidated memories can be disrupted illustrating that even long-term memory traces were not fixed in time (Misanin et al., 1968; Schneider and Sherman, 1968). The ever-changing nature of long-term memory traces has been more intensively investigated since 2000 in light of reactivation manipulations. Throughout these studies, the main focus was made on memories in their active form but little is known on the inactive state of the memories. Therefore, many questions remain to be answered. For example, can memories survive without any subsequent reactivation? Or more precisely, what is the longest retention interval during which a memory can be maintained prior to reactivation? And what is the minimal number of exposures that would allow memory maintenance?

In that context, the aim of my thesis was to focus on these memories that were not activated for a long time. This original work was developed under the supervision of Simon Thorpe as part of a scientific project entitled: Memory Mechanisms in Man and Machine (see the webpage: http://m4.ups-tlse.fr/?page_id=57) and financed by a European Research Council grant.

I developed my thesis around two main axes: a theoretical approach that constituted the first and second chapter of this manuscript and an experimental framework based on behavioral (chapters 3 and 4) and neuroimaging data (chapters 5, 6, 7).

1.2 Substantial progress

The theoretical framework that we developed allowed us to put the notion of inactive memories into a dynamic model of inferences that is common to perceptual and memory processes.

According to the two behavioral experiments that we carried out we can claim that:

“Three repetitions are enough to create a memory trace explicitly accessible at least a decade later and without reactivating the information in the intervening period”

Let us dissect and consider the important parts of this sentence:

“three repetitions”: Under this threshold it seems that memories might be more easily forgotten.

“explicitly accessible”: this long-lasting memory trace can be the basis of voluntary actions when individuals are directly involved in the retrieval of their memory.

“at least a decade later”: And potentially for a lifetime...

“without reactivating the information”: The availability of a memory trace does not seem to rely on reactivations -even scattered in time- after the last exposure.

Importantly, the reactivation of these dormant memories might be identified on the EEG signal by using decoding approaches based on single-trial analysis.

Unlike many EEG studies, we did not investigate long-term retrieval by considering participants' subjective state (retrieved/forgotten...) but according to the retention interval between learning and retrieval, a variable that we could manipulate in the paradigm and therefore directly study (VanRullen, 2011).

2. Theoretical considerations

Several questions might arise from the claim above. Indeed to get the message clearer in the previous chapters, we did not spend much time considering some theoretical aspects that were involved. Now it is time to look more closely at some critical points.

2.1 Identification, familiarity, novelty

2.1.1 Identification in a forced-choice task

In chapter 1 we presented two criteria often used to encompass the large variety of retrieval processes: the “familiarity” and the “knowing” criterion. The “knowing” criterion was defined as an identification criterion where a specific hypothesis was selected from the set of all possible hypotheses proposed. The familiarity criterion is less conservative and simply separates all the hypotheses of a given task in 2 classes: novel and familiar. The boundary is made from the posterior probability that the class “novel” would be more likely and as such, the “familiarity” criterion could be also called “novelty” criterion. In a forced-choice task the classes: “novel” and “familiar” are reduced to two hypotheses corresponding to the two items proposed. This means that the distinction familiar/novel merely relies on a “knowing”

criterion which is an identification criterion. Correct responses in a forced-choice task might therefore not necessarily rely on familiarity/novelty processes. Indeed, we observed that most of the participants who could correctly identify simple drawings in a forced-choice task (chapter 4) could not report any feeling of either familiarity or novelty. We suggest that the Yes/No task in which differentiation is made based on a “familiarity” criterion would be a better design to bring into light familiarity and novelty judgments.

2.1.2 Familiarity versus novelty

In a design that constrains participants to use a “familiarity” criterion to respond (such as in a Yes/No task), responses can be based on either a familiarity or a novelty judgment or even both. As for any class or hypothesis, we suggest that familiarity and novelty judgment are the result of evidence accumulation. In such perspective, familiarity and novelty are not necessarily symmetrical responses as suggested by recent experimental findings (Barbeau et al., 2017). The data in chapters 6 and 7 are particularly demonstrative: participants are good to identify new pictures in a cued-recall task, but it is more difficult for them to rely on a familiarity judgement.

2.2 Consciousness and memory

2.2.1 Difference between implicit/explicit memory and conscious/unconscious retrieval

In this manuscript we particularly focused on explicit memories, the ability to voluntarily retrieve some memory content. It is different from implicit memory where retrieval is not voluntary and can be inferred from behavioral performance. It is important to make the difference between implicit/explicit and conscious/unconscious processes. Particularly, in chapters 3 and 4 we found that participants involved in an explicit retrieval task could be conscious of their memory or not. While no consensus holds for the definition of consciousness, it is generally assumed that it implies a second order judgement (Rosenthal, 1986): for example: “I know that I recalled the correct name of the TV program”. In that sense consciousness binds knowledge and metaknowledge together (Koriat, 2007). As specified in chapter 1, the report of confidence level is a good way to assess metamemories and metaknowledge in general.

2.2.2 Unconscious reactivations

In this thesis we showed that participants could retrieve explicit memories for stimuli that were not presented again since the time of exposure. However it might be possible that these memories could have been reactivated by internal or external cues. It is hard to fully exclude such possibility for TV programs that are part of a rich network of associations. However this seems less likely for simple drawings such as the ones we used in chapter 4. In addition, a third of these pictures were abstract stimuli that were not easy to verbalize and did not resemble any particular object from our environment. Interestingly, participants' performance was similar between these stimuli and more concrete drawings. This confirms that explicit reactivations of the memory trace might be not necessary to maintain a memory trace, as already found in long-term memory for abstract Gaussian noise (Agus et al., 2010; Viswanathan et al., 2016). However, and even using meaningless stimuli, it is still possible that long-term maintenance would rely on unconscious reactivations of the memory trace. Unconscious reactivations could emerge from the spontaneous firing of neurons that code for a specific memory but this signal might be too weak to create a subjective feeling of retrieval. To what extent this physiological property would be linked to very long-term maintenance remains to be tested.

2.3 Concepts and neuronal selectivity

2.3.1 Selectivity to concepts

In this manuscript we hypothesized that active memories might rely on highly excitable neurons that can be easily reactivated whereas inactive memories would be maintained via silent neurons. According to our view, this implies neurons to be highly selective to a particular stimulus and therefore might concern neocortical neurons. But what does it mean to be selective when considering content-rich stimuli such as a TV programs like 'Friends'? Presumably, a neuron selective to 'Friends' won't keep firing all along the TV series... Some neurons would fire specifically for a given character, some others when the beginning of the sound of the TV theme is played, so on and so forth. Neurons would therefore encode specific concepts. But is not 'Friends' a concept in itself? So, in theory, if a neuron is specific to a TV program it might fire all along the audiovisual content. There is evidence that a neuron in the human entorhinal cortex selective to 'The Simpsons' fired throughout the 5-seconds of the TV series clip. Even more remarkably, the same neuron also fired in the absence of external cues, when the same individual was recalling having seen this clip previously and until he could

think of another clip (Gelbard-Sagiv et al., 2008). Of course, 5-seconds are nothing compared to the whole duration of a movie and it might be interesting to see whether some neurons could fire for longer periods. This sustained activity might correspond to the maintenance of the concept in working memory, the first step before long-term storage.

2.3.2 Selectivity through temporal contiguity

Sustained delay activity was found in IT neurons in monkeys (Miyashita, 1988; Sakai and Miyashita, 1991). Interestingly, data showed that if an image was presented in a sequence with other pictures, the neuron selective to the first image would enlarge its selectivity to the following six pictures presented (Miyashita, 1988). This effect would be the largest for the closest picture and would decrease over the number of pictures. Sustained delay activity could then allow the association between successive stimuli inside the same neuron. This widening of neuronal selectivity to neighboring items does not only concern pictures following the first image but also occur for preceding items. Indeed, recent data in the human MTL showed that a neuron selective to a picture can start firing previous to its presentation if it is embedded in a given fixed sequence (Reddy et al., 2015). Overall, the association of stimuli through temporal contiguity would be a simple mechanism by which different views of the same stimulus or different items belonging to the same stimulus would be associated together to create a specific concept and would not require complicated computations (Miyashita, 1993).

2.3.3 Flexible selectivity for creating associations

The “anticipatory” effect described just before in Reddy et al. appeared after only 11 trials, showing the large flexibility of the MTL structures for creating new associations in individual neurons. This flexibility can happen even faster in the absence of a temporal delay between two stimuli. This was recently shown that in the human MTL where neurons that were at first only selective to a specific stimulus (such as a patient’s relative) could become selective to another stimulus (such as the Eiffel Tower) within a few trials if the two were presented simultaneously (Ison et al., 2015). It would be interesting to test how this selectivity would evolve over a large number of trials. According to our view, repetition might sharpen the neuronal population selective to a specific a memory. But as mentioned in chapter 2, other theories suggest a decrease of neuronal activity over the repetitions. During my thesis, I had the opportunity to design and carry out an experiment that tried to answer this question. I presented new associations of stimuli to epileptic patients in Toulouse for whom it was necessary to implant electrodes to localize the epileptic region. In this experiment, the patients

had to learn three associations made of a name (a sound) and a picture. During many test trials patients were tested on their ability to recognize from two pictures the one corresponding to the name heard. With this design, we were able to record about 400 trials over several days of recording. This large number of trials might give us a better insight concerning the effect of repetition on neuronal selectivity. The data still needs to be analyzed.

2.3.4 Associations in the neocortex

The MTL holds a key role in the binding of items and concepts. Another critical region that deserves attention and especially when considering long-term memory is the prefrontal cortex (Euston et al., 2012). The dorsolateral PFC for example is particularly involved in building relationships between items such as pairs of words (Blumenfeld and Parks, 2011). As for the MTL, neurons in the PFC show highly sustained activity that allows the maintenance of related items in working memory (for early evidence for e.g: (Funahashi et al., 1989; Fuster and Alexander, 1971). Interestingly, neurons in the PFC are strongly selective to particular concepts such as faces (Scalaidhe et al., 1999) or can link an object within its location (Rainer et al., 1998).

Most of the neurons recorded in the PFC show categorical selectivity driven by a given task and is more referred to a relative rather than an absolute selectivity (Duncan, 2001). For example, a neuron that is found selective to a dog but not to a cat in a given set of stimuli, would then become selective to a car but not to another type of car in another set of stimuli (Cromer et al., 2010). To be effective, the concepts encoded in one neuron should have distinct physical representations -such as dog and car- to avoid any competition between the items. Instead, these associations would reflect the learning of abstract rules that goes beyond physical properties (Wallis et al., 2001).

2.4 Forgetting

“Mental states of every kind, -- sensations, feelings, ideas, -- which were at one time present in consciousness and then have disappeared from it, have not with their disappearance absolutely ceased to exist. Although the inwardly-turned look may no longer be able to find them, nevertheless they have not been utterly destroyed and annulled, but in a certain manner they continue to exist, stored up, so to speak, in the memory. We cannot, of course, directly observe their present existence, but it is revealed by the effects which come to our knowledge with a certainty like that with which we infer the existence of the stars below the horizon.” (Ebbinghaus, 1885).

2.4.1 The importance of forgetting

Whatever it is, a loss of access or loss of information forgetting seems to play an important role for normal brain functioning. Indeed we usually complain about forgetting in our daily life but it helps us not endure the burden of being taken back every time to our past. The ability to remember everything from our past was nicely illustrated through a fictional character, Funes in the book: *Funes the Memorious* written by Jorge Luis Borges in the mid 50's. It turned out that such story could in fact be true. The first case reported, was A.J., a woman in her late 30s who could remember every day of her life since her teens (Parker et al., 2006). In the letter she wrote to explain her case, she herself reported: "Most have called it a gift but I call it a burden. I run my entire life through my head every day and it drives me crazy!!!..." Following this case, other individuals were found with such extraordinary ability called Highly Superior Autobiographical Memory (LePort et al., 2012). This ability seems to be involuntary and is not deliberately triggered as for memory champions who use voluntarily different strategies to increase the retrieval of semantic memories (Parker et al., 2006). In particular, an excessive development of MTL structures would explain higher autobiographical abilities (LePort et al., 2012).

As we see, forgetting might prevent the constant revival of old memories. It could be trace-dependent meaning that the content available in memory would change over time but also cue-dependent where the pieces of information present in our environment would not be specific enough to re-access the memory. Rather than keeping them separate, the combination of these two hypotheses might explain retrieval performance in a given task (Tulving and Madigan, 1970).

2.4.2 The changing of a memory trace

Since the beginning of the 20th century two main factors have been proposed to account for forgetting: interference and decay (for a review see (Wixted, 2004)). Interference would occur during the memorization of competing stimuli. If some new information is learnt and, if it is similar enough to previous memories, the old information might be forgotten (retroactive interference). Conversely, the presence of old memories might impair the retrieval of similar and new memories (proactive interference). Unlike interference, decay would not assume any competition between similar items, but would explain forgetting as the consequence of the passage of time. Recently it has been proposed that both interference and decay might account for forgetting but that they would involve different brain structures (Hardt et al., 2013; Sadeh

et al., 2014). Decay would explain forgetting in the hippocampus, whereas interference would be found mainly in neocortical structures.

- Decay in the hippocampus

Via efficient pattern separation, the hippocampus is able to create distinct neuronal representations for many events of our life (Rolls, 2016). Memories stored in the hippocampus are therefore preserved from potential interference phenomena: their neuronal representations are too different to lead to some competition. It has been shown that pattern separation mainly occurs in the dentate gyrus of the hippocampus, a region that supports neurogenesis throughout the all individuals' life. The side effect of neurogenesis is that it might induce the decay of old memories that are not used (Frankland et al., 2013). Episodic memories that strongly rely on hippocampal functions might be potentially erased from the hippocampus or re-written. This would explain why it would be difficult for episodic memories to survive over time in the absence of reactivations (Sadeh et al., 2014) and why none of our participants could retrieve episodic memories after about 10 years.

- Interference in the (neo)cortex

Episodic memories tend not to resist very well over time but semantic memories and familiarity seem more robust. In the same way as pattern separation is a key mechanism in the hippocampus, pattern categorization might be performed by (neo)cortical structures (Rolls, 2016): the prefrontal cortex as well as cortical regions from the MTL for semantic memories (Binder et al., 2009) and only cortical regions of the MTL for familiarity (Sadeh et al., 2014). The general idea would be to cluster similar neuronal representations and make a generic and sparse representation of them. While sparse coding would constitute an efficient way for storing information in (very) long-term memory, the neuronal coding might be more prone to interference, due to some invariance in the neuronal representation (Sadeh et al., 2014). However our data show that 12 years later participants were still able to differentiate similar pictures demonstrating the high specificity (low invariance) of the patterns stored.

2.4.3 The role of the cues

In this thesis we showed that memories that were thought to be lost can be retrieved years later. In this case, forgetting was due to a failure to access the information rather than to a loss of information. The importance of cues is then crucial to bring memories back to mind (Habib and Nyberg, 2008; Tulving, 1974). But if the cues were efficient enough to retrieve semantic

or perceptual memories and to generate some familiarity, they failed to revive every contextual detail that was present at the moment of the encoding. This observation is in agreement with the literature showing that it becomes more difficult to retrieve episodic memories as time elapses.

In the chapter 1 of this thesis we argued that the retrieval of episodic memories might be on a continuum with the retrieval of semantic information where secondary pieces of evidence would be gathered to re-create the original and unique episode at the time of encoding. However, cues that are generally provided in experimental designs are directed to the main perceptual or semantic information and give no information about the “secondary” details that would make a memory different from another one. For example, we directly showed audiovisual excerpts of TV programs (a cue) to elicit the retrieval of their titles (semantic information). But we did not/could not provide any contextual cues for a specific event that happened one day during the watching of an episode which would have helped the recollection of this particular episode. If such contextual details, unique to a particular episode, would have emerged, they would have been elicited internally by the individuals. In that sense it would be the same as a free recall when considering semantic information.

More importantly, while perceptual and semantic memories are tested with targets that present some differences to avoid potential interferences, the context in which they are presented, - usually a “boring” room in the lab- does not change much. This is particularly the case for our study about long-term memory for drawings. How much did contextual information differ from the drawing presented on the first page of the binder to the 48th one? So, how could participants revive the particular moment when they saw a specific drawing? From all of the reports given by the participants, seven of them were spontaneously described as “flashback” with six of them for pictures correctly identified. It seems then that participants would still have the ability to mentally travel back to the time of the experience. But in the conditions of our study, we did not have a better option than to consider these responses as the result of familiarity processes. However in refined designs, it is not unlikely that episodic memory could have emerged from such strong feeling.

According to these two arguments: 1/the absence of contextual cue during retrieval and 2/ the absence of varied contextual details during memory encoding, it is yet not clear whether episodic memories would be as easy to forget as we suspect.

Therefore how could we test (very) long term episodic memory? As a *mise en abyme* (a story in a story) we could make participants watch audiovisual content and test them years later. As already done for a retention interval up to 9 months (Furman et al., 2007), the idea would be to present snapshots of a narrative that was presented years ago and ask them details about what happened after (or before) the still frame (cued- recall). If too, difficult they might have the option of watching two short clips and would be asked to determine which of the two was already presented to them (recognition) (Tang et al., 2016). Along these lines, Meyi Dulème, extended her M2 internship in the lab and ran a pilot experiment where she asked the participants who watched 480 audiovisual clips, 3s-long, to come back to the lab five months later. A part of these results was presented at the end of chapter 7. Interestingly, participants were still 70% correct at identifying an old clip from a new one when they were replayed.

Even more realistic, the use of still snapshots or short audiovisual clips from real-life experience might be valuable in the retrieval of episodic memories. SenseCams (Hodges et al., 2006) were developed to that purpose. They are small wearable cameras that can take pictures of individual's daily life, passively, what is commonly referred as lifelogging. In research, the use of this tool was motivated at first for clinical reasons in patients with autobiographical memory disabilities. In light of the positive results obtained (e.g (Berry et al., 2009; Loveday and Conway, 2011) it deserves more importance in cognition in general (Doherty et al., 2012).

2.5 False memories and confabulations

Along with forgetting, another main point is to assess and control for the truth of the memories retrieved. Far from being perfect video tapes that record and replay every event the way they originally happened, our memories are the result of reconstruction processes. As for perception, memories are prone to illusions, the two being tightly linked (Roediger, 1996). No doubt, our memories are fallible, malleable and even false or fabricated (Loftus, 2005). There are apparently no exceptions to this rule, no matter what age or whether we show Highly Superior Autobiographical memory or not (Patihis et al., 2013). False memories about semantic information can be formed within only four seconds after encoding (Atkins and Reuter-Lorenz, 2008) and might be due to associative inferences where novel information can be directly linked to previous knowledge without the two being experienced together (Carpenter and Schacter, 2016; Zeithamova et al., 2012). Regarding autobiographical memories, false memories might be created to ensure consistency with our self and would involve a tradeoff between concordance with past event (likelihood) and coherence with our

self (prior about the self) (Conway and Loveday, 2015). According to the authors, most of autobiographical memories would be a combination of low-concordance (again, according to the way it was studied but it might be higher) and high coherence. In addition to low-concordance, when coherence is disrupted it might result in confabulations.

Recollective confabulation is defined as the retrieval of specifics from an experience that never happened which is associated with a strong feeling of remembering (Moulin et al., 2005). This is close to *déjà vu* reported in the healthy population, the impression of having already experienced a situation associated with the awareness that the actual experience should be considered as new (O'Connor and Moulin, 2010). Recollective confabulation is a constant phenomenon that impairs the daily life of the patients suffering from this dementia who see no interest in new personal events thought to be already experienced (O'Connor et al., 2010). A striking feature of such dementia is that every detail from an actual experience is reported as being familiar by the patients (Moulin, 2013), a judgement that should never happen considering that every single experience is a mixture of already known or familiar items and new ones (Tulving and Donaldson, 1972). According to the evidence accumulation model we developed in the Introduction of the thesis, the amount of evidence accumulated in parallel for the relevant features of an actual scene might fall in beyond the familiarity and knowing criterion but it will not be the case for other features. In the case of *déjà vu* there might be no room left for novelty detection and the amount of evidence of all features of the scene might reach the two criteria. According to this view, whereas in patients suffering from Alzheimer's disease the familiarity criterion might be too conservative and might result in a bias toward novelty, the opposite might account for recollective confabulation. In the latter case, a permanent bias toward familiarity might arise from the presence of a very lenient familiarity criterion or its complete absence. This view is in line with recent interpretations from pathological cases that consider that recollective confabulations might be due to false familiarity judgment (Moulin, 2013) and that the entorhinal cortex might account for such impairment (O'Connor et al., 2010).

But as pointed out by Loftus, "although experimental work on the creation of false memories may raise doubt about the validity of long-buried memories, [...] it in no way disproves them"(Loftus, 1997). This gives more credit to the reports obtained from the participants who were tested on the memory of old TV programs. But indeed, even for the two participants

who could spontaneously retrieve the title of a TV program, the contextual details that were reported were not perfectly exact.

2.6 The role of affects

Subjective affect is a generic term that we will use in this section. It includes emotions defined as a short-lasting affective state in reaction to internal or external stimuli (Scherer, 2009) and mood which is longer lasting (few minutes) and might be more an intrinsic feature (Bower and Forgas, 2000). In this thesis we were particularly interested in testing long-term retrieval for stimuli that were fairly neutral and that would not particularly arouse strong feelings. Indeed, such memories would have been more prone to be spontaneously reactivated after the period of exposure, something that we wanted to avoid as much as possible. This effect can be explained by considering subjective affects as very good cues to re-access past information (Buchanan, 2007). It might account for mood-dependent and mood-congruent memories where the retrieval of memories might be influenced by the affective state of individuals. In conditions where subjective affect was not supposed to play a predominant role, we found that three repetitions might be a sort of cutoff value from which we could expect very long-term retrieval. However we might expect that for events associated with strong affects, one exposure might be enough to be retained for a lifetime. This is the case of flashbulb memories, memories that are associated with strong vividness and details about striking events such as the assassination of John Fitzgerald Kennedy or the 9/11 day. But again, because of the shock they provoked, many reactivations would have made it easier long-term maintenance of these memories (Neisser, 1997). In other cases, the emotional charge is so high that individuals are not able to cope with their real-life experience and suffer from post-traumatic stress disorder.

3. Future directions

3.1 Limits

3.1.1 Replication and extension

The claim that we make: *“Three repetitions are enough to create a memory trace explicitly accessible at least a decade later and without reactivating the information in the interleaving period”* comes from behavioral data resulting from two experiments that were conducted using two different paradigms. Further studies should confirm the results obtained.

We suggest that the experiment concerning the memory of old French TV programs might be continued and extended to individuals from all over France. Results would be easily collected via a web platform. The same type of experiment could also be performed in other countries using their own audiovisual archives.

Furthermore, the combination of behavioral performance together with neuroimaging information might provide valuable information. The memory of the pictures presented in the DMS-48 could be assessed while measuring the EEG activity of participants that have been not re-tested yet. Again, other stimuli could be used in other labs as long as the material was not rehearsed for several years.

3.2 Linking behavior and decoding analyses

We showed that the difference of brain processing between stimuli seen 3 weeks ago and new stimuli could be decoded based on the EEG analysis of single trials. We hope that this is going to be a first step and that further studies would be done on longer retention intervals between the encoding and the testing phase. In chapters 6 and 7, behavioral and EEG data were analyzed separately. However, more information might be gathered if the two would be combined together. To manipulate data that is not expressed in the same measure unit it has been recently proposed to use representational dissimilarity matrices (RDMs) (Kriegeskorte et al., 2008). Each row and column of the matrix corresponds to a specific experimental condition and each value inside the matrix corresponds to the dissimilarity (1-correlation) between two conditions in a given measure unit. This enables the creation of RDMs for every type of information gathered in an experimental set up -from behavioral or EEG data- that can be correlated together to get the overall difference between two conditions (Kriegeskorte and Kievit, 2013).

3.3 From Bayesian to integrated information

In chapter 2 we showed that the Bayesian theory could be considered in a hierarchical architecture where the a posteriori distribution generated by neurons in a specific layer would correspond to the a priori information in the next layer. In this section, we will discuss the potential emergence of a conceptual structure from this physical architecture. Every time, the new shaping of this conceptual space would lead to a unique phenomenological experience.

3.3.1 Information based on conditional entropy

As explained by Shannon (Shannon, 1948), information corresponds to the reduction of uncertainty. Uncertainty is a direct measure of the variability of a probability distribution (H_i). The more there are possibilities, the more there is variability and the larger is the entropy of the distribution $S(H_i)$.

In Bayesian terms, the maximum uncertainty corresponds to the set of all the different possible hypotheses ($H = \Sigma H_i$) and refers to a repertoire of prior probabilities. This “a priori repertoire” corresponds to the maximum entropy and can be considered as the overall amount of available information ($S(H)$). By taking into account an actual event (E) this “a priori repertoire” reduces to a more certain subset of probabilities ($H|E$) that are referred to as “a posteriori repertoire”. Because the variability of the “a posteriori” distribution is smaller, its entropy is reduced: $S(H|E)$ and corresponds to the amount of accessible information. The quantity of information generated is the “difference” between the maximum entropy $S(H)$ and the conditional entropy (also called equivocation) $S(H|E)$ and can be measured in bits (if the logarithm used is a basis of 2).

3.3.2 Integrated information

If the conditional entropy ($S(H|E)$) is merely the product of the entropy of each hypothesis in the “a posteriori” repertoire, this means that the information generated was not integrated beyond its parts (Tononi, 2008). Therefore, integrated information corresponds to the relative entropy between the conditional entropy of the whole repertoire $S(H|E)$ and the sum of the entropy for each hypothesis. Information that emerges from such integration is therefore indivisible. It is an entity that would fall apart if one wants to separate it into smaller pieces. The set of elements from which integrated information emerges is called a complex and can simply correspond to a set of neurons (Balduzzi and Tononi, 2009). This complex would correspond to a given concept and can be viewed as a point in a multidimensional-space where each axis corresponds to a hypothesis among all the possible a priori hypotheses. This point/concept defines a particular shape or quale in that space. As such the increase of integrated information in a given system would create a more refined shape which might be the basis of every subjective experience (Oizumi et al., 2014).

3.3.3 Our memories into quales

Our ability to re-experience past events or to predict future situations would be then directly tied to the formation of specific quales in our brain. The sharpness and distinctiveness of each quale would be an indicator of the accuracy of our mental images, in the past or in the future. In this mental traveling, the reinstatement of specific episodes of our life would be similar to the creation process of perfect diamonds in a jeweler's hands. Such expertise would not be reached every time and diamonds might give place to simpler gems in familiar occasions or to shapeless rocks in more uncertain conditions.

4. Conclusion

4.1 A personal view to my thesis

Shapes do not appear or re-appear instantaneously but instead, reflect a cascade of brain processes that gradually refine the representation over time (chapter 1). Such representations are fleeting and might quickly fly away to leave some room to constantly new original experiences except if we use the right trick to bring them back (chapter 2). As for mountain peaks, the shape of our memories are intended to evolve. The repetition of similar patterns of neuronal activation will lead to increasingly more sharpened summits clearly outlined against the sky and easily visible from miles away (chapter 7). In parallel with strong and local activities, more subtle and broader activations will also reshape the overall landscape to firmly hold the entire structure on strong foundations. However, the termination of a specific neuronal activity will show the beginning of erosion, the highest peaks being affected first, before the gradual decline of the whole formation (chapter 6). Sharp and high mountains are now rounded hills and broad valleys that still hold some pieces of information but need adapted tools to dig them up (chapters 3 and 4).

4.2 Can one see the shape of our memories?

In fairness, my main motivation for doing my PhD was to better understand the way subjective experience emerges from neuronal activity. Simon gave me the opportunity and the financial support to tackle this issue on the side of memory. A great deal when one knows that memory encompasses all the variety of our subjective experiences! With Chris we showed that the quality of our subjective experiences is the result of a building of evidence over time. I then refined this framework for the specific case of dormant memories, the main focus of my thesis. With all the collaborators involved in the project we ended up discovering that memories might be more robust to decay than previously thought. For example, who could

have imagined that the memory of simpler pictures could still drive explicit behaviors 10 years later? But of course, the passage of time and potential interferences might have altered the accuracy of our memories. We proposed to use decoding analyses on the EEG signal to show these subtle changes. However decoding does not provide enough information concerning the quality of the information generated and for now the shapes of our memories has not been revealed.

After all this work a main question remains unanswered to me: from neurons that create the biological structure to our memories, what is the physical support to our mental images? If they are “shapes” then: “can we see the shapes of our memories?”

OPENING. Beyond matter, from Alan Turing to Cymatics

“Witnessing the processes that bring [...] forms into being, offers us penetrating insight into how an elaborate web of vibrations interacts to create the world we perceive” (Volk, 2010).

After spending a lot of effort and money in running thousands of thousands of experiments with the best technology ever, we don't know yet how our subjective experiences emerge in the brain. Of course we know that neurons, our brain cells, play a big role and somehow constitute the biological substrate of that information but for now, brain's code hasn't been cracked.

1. Standing waves

In 1952, Alan Turing wrote a paper called *The Chemical Basis of Morphogenesis* (Turing, 2007), in which he describes the role of morphogens in the creation of self-regulated patterns such as the embryo's form (see also Kondo et Miura, Science, Review, 2010). This theory is known as reaction-diffusion or Turing model and relates to morphogens that are small molecules like peptides. By interacting within each other through chemical reactions and diffusing into a system such as a tissue, morphogens allow the creation of different steady states. Among them, stationary patterns can emerge which are characterized by stationary waves with finite wavelength. As their names suggest, stationary or standing waves do not propagate and are the result of the interference between two waves that travel in opposite directions. These types of waves, also known as Turing patterns, are able to generate a large variety of spatial patterns that are independent of any preexisting positional information (prepattern) like the disposition of zebra stripes or feathers in chick (Jung et al., 1998). More interestingly this does not only apply to morphogens but for any system satisfying the two conditions: short-range positive feedback with a long-range negative feedback as shown by Meinhardt and Gierer (Meinhardt and Gierer, 2000) and including neuronal interactions (Swindale, 1980). In the latter case, they are the results of synchronous neuronal activity (Benucci et al., 2007).

The interaction between neurons is able to generate standing wave patterns, the same way as for morphogens. And when expressed in a given environment, it induces the creation of

different spatial patterns that are independent of the neuronal assembly configuration also called pre-pattern. Neuronal information/brain's code would lie on that new pattern that needs a physical support to be perceived.

2. Cymatics

The term cymatics comes from the Greek root kyma: *wave* and was first introduced by Hans Jenny (1904-1972) whose work was in straight continuation with the German physicist Chladni. In his book *Entdeckungen uber die Theorie des Klanges* (“Discoveries in the Theory of Sound”) published in 1787, Chladni showed that by drawing a bow over a steel plate sprinkled with sand, the vibration induced in the plate varies from a region to another which makes the sand migrates from the areas of strong to low intensities. This resulted in the formation of distinctive patterns; the so-called Chladni's figures (Figure O.1)

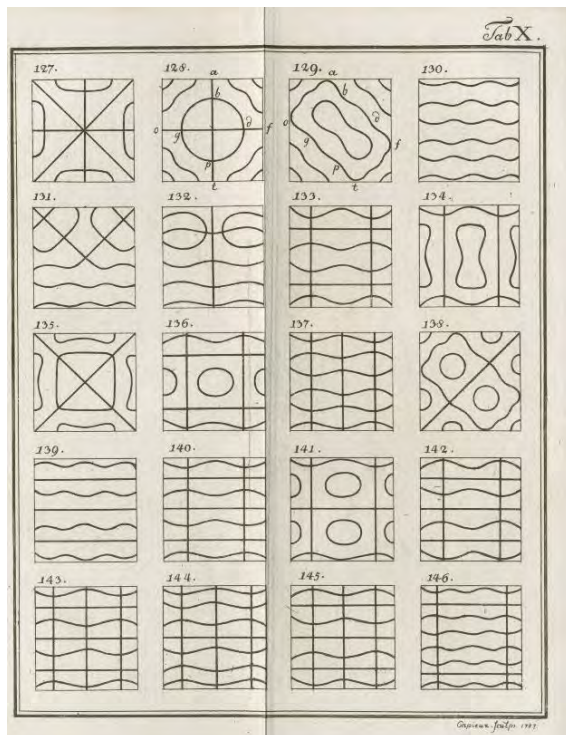


Figure O.1 Exemplar of Chladni's figures (from (Chladni, 1787))

Cymatics offers the great possibility to turn sounds and more generally any kind of vibrations into specific forms, a rare glimpse into the hidden dynamics of nature (Volk, 2010). Information, “in” – “formation”, literally means giving form and a shape constitutes its material expression.

3. Harmonics in the brain

Constantly, the brain receives vibrations from what constitute our external world. By exciting neuronal populations, these vibrations coming from the outside world would create brain oscillations. Brain oscillations then would convey information that can be frozen into a new space-time dimension if played into a Chladni's plate. These vibrations would be transformed into patterns to recreate the world in a certain way. Shapes created on the plate are just the expression of the interplay between a wave and any other support like a plate. As a result, the plate freezes the wave at different time points that depend on its own properties. Importantly, the generated shapes directly depend on the type of medium used: water, plate... and on its physical properties such as thickness or diameter. **This means that the information we receive and by extent, the reality that we perceive, do not exist per se, but instead is the expression of the interplay between two entities.** Chaldni's figures coupled with neuroimaging techniques might help catching information generated by the brain and might be a key to apprehend consciousness. A first step has already been done successfully in this direction (Atasoy et al., 2016).

FINANCIAL SUPPORT

I was very lucky to have the financial support that I needed to perform all my studies. This support was provided by two funding agencies: the European Union's Seventh Framework Programme (FP7/2007-2013)/ERC grant agreement n°323711-Project M4) and the French Agence Nationale de la Recherche (ANR)/ANR-12-BSH2-0010-Project ELMA). I want to acknowledge these two funding agencies for their substantial financial contribution.

REFERENCES

- Abeles, M., Vaadia, E., Bergman, H., 1990. Firing patterns of single units in the prefrontal cortex and neural network models. *Comput. Neural Syst.* 1, 13–25. doi:10.1088/0954-898X
- Acunzo, D.J., Mackenzie, G., van Rossum, M.C.W., 2012. Systematic biases in early ERP and ERF components as a result of high-pass filtering. *J. Neurosci. Methods* 209, 212–8. doi:10.1016/j.jneumeth.2012.06.011
- Adolphs, R., Damasio, H., Tranel, D., Damasio, A.R., 1996. Cortical Systems for the Recognition of Emotion in Facial Expressions. *J. Neurosci.* 16, 7678–7687.
- Aggleton, J.P., Brown, M.W., 1999. Episodic memory, amnesia, and the hippocampal–anterior thalamic axis. *Behav. Brain Sci.* 22, 425–489. doi:10.1017/S0140525X99002034
- Agus, T.R., Thorpe, S.J., Pressnitzer, D., 2010. Rapid Formation of Robust Auditory Memories: Insights from Noise. *Neuron* 66, 610–618. doi:10.1016/j.neuron.2010.04.014
- Alberini, C.M., 2005. Mechanisms of memory stabilization: are consolidation and reconsolidation similar or distinct processes? *Trends Neurosci.* 28, 51–6. doi:10.1016/j.tins.2004.11.001
- Allen, E.A., Erhardt, E.B., Calhoun, V.D., 2012. Data visualization in the neurosciences: overcoming the curse of dimensionality. *Neuron* 74, 603–8. doi:10.1016/j.neuron.2012.05.001
- Anderson, B., Mruczek, R.E.B., Kawasaki, K., Sheinberg, D., 2008. Effects of familiarity on neural activity in monkey inferior temporal lobe. *Cereb. cortex* 18, 2540–52. doi:10.1093/cercor/bhn015
- Anderson, J.R., 1995. A simple theory of complex cognition. *Am. Psychol.* 51, 355–365.
- Anderson, J.R., Matessa, M., 2003. The rational analysis of categorization and the ACT-R architecture, in: Oaksford, M., Chater, N. (Eds.), *Rational Models of Cognition*. pp. 197–217.
- Anderson, J.R., Schooler, L.J., 1991. Reflections of the environment in memory. *Psychol. Sci.* 2, 396–408.
- Andrillon, T., Kouider, S., Pressnitzer, D., Agus, T., 2015. Perceptual Learning of Acoustic Noise Generates Memory-Evoked Potentials. *Curr. Biol.* 25, 1–7. doi:10.1016/j.cub.2015.09.027
- Atasoy, S., Donnelly, I., Pearson, J., 2016. Human brain networks function in connectome-specific harmonic waves. *Nat. Commun.* 7, 10340. doi:10.1038/ncomms10340
- Atkins, A.S., Reuter-Lorenz, P.A., 2008. False working memories? Semantic distortion in a mere 4 seconds. *Mem. Cognit.* 36, 74–81. doi:10.3758/MC.36.1.74

- Baars, B.J., 1988. A cognitive theory of consciousness. NY: Cambridge University Press.
- Bacon-Macé, N., Macé, M.J.-M., Fabre-Thorpe, M., Thorpe, S.J., 2005. The time course of visual processing: backward masking and natural scene categorisation. *Vision Res.* 45, 1459–69. doi:10.1016/j.visres.2005.01.004
- Bahrnick, H.P., 1984. Semantic Memory Content in Permastore: Fifty Years of Memory for Spanish Learned in School. *J. Exp. Psychol.* 113, 1–29.
- Bahrnick, H.P., Bahrnick, P.O., Wittlinger, R.P., 1975. Fifty years of memory for names and faces: A cross-sectional approach. *J. Exp. Psychol. Gen.* 104, 54–75. doi:10.1037//0096-3445.104.1.54
- Balduzzi, D., Tononi, G., 2009. Qualia: the geometry of integrated information. *PLoS Comput. Biol.* 5, e1000462. doi:10.1371/journal.pcbi.1000462
- Barbeau, E., Didic, M., Tramon, E., Felician, O., Joubert, S., Sontheimer, A., Ceccaldi, M., Poncet, M., 2004. Evaluation of visual recognition memory in MCI patients. *Neurology* 62, 1317–1322. doi:10.1212/01.WNL.0000120548.24298.DB
- Barbeau, E., Chauvel, P., Moulin, C.J., Liégeois-Chauvel, J.R., Liégeois-Chauvel, C., 2017. Hippocampus duality: Memory and novelty detection are subserved by distinct mechanisms. *Hippocampus* 27, 405–416. doi:10.1002/hipo.22699
- Barbeau, E.J., Pariente, J., Felician, O., Puel, M., 2011. Visual recognition memory: a double anatomo-functional dissociation. *Hippocampus* 21, 929–34. doi:10.1002/hipo.20848
- Barlow, H., 1994. The neuron doctrine in perception, in: Gazzaniga, M. (Ed.), *The Cognitive Neurosciences*. Boston; MIT Press, pp. 415–435.
- Barlow, H., 1972. Single units and sensation: a neuron doctrine for perceptual psychology? *Percept* 1, 371–394.
- Barragan-Jason, G., Besson, G., Ceccaldi, M., Barbeau, E.J., 2013. Fast and Famous: Looking for the Fastest Speed at Which a Face Can be Recognized. *Front. Psychol.* 4, 100. doi:10.3389/fpsyg.2013.00100
- Barron, H.C., Vogels, T.P., Emir, U.E., Makin, T.R., O'Shea, J., Clare, S., Jbabdi, S., Dolan, R.J., Behrens, T.E., 2016. Unmasking Latent Inhibitory Connections in Human Cortex to Reveal Dormant Cortical Memories. *Neuron* 90, 191–203. doi:10.1016/j.neuron.2016.02.031
- Bayley, P., Wixted, J., 2008. Yes/no recognition, forced-choice recognition, and the human hippocampus. *J. Cogn. Neurosci.* 20, 505–512.
- Bayley, P.J., Hopkins, R.O., Squire, L.R., 2006. The Fate of Old Memories after Medial Temporal Lobe Damage. *J. Comp. Neurol.* 26, 13311–13317. doi:10.1523/JNEUROSCI.4262-06.2006
- Bayley, P.J., Hopkins, R.O., Squire, L.R., 2003. Successful Recollection of Remote Autobiographical Memories by Amnesic Patients with Medial Temporal Lobe Lesions. *Neuron* 38, 135–144.

- Benchenane, K., Peyrache, A., Khamassi, M., Tierney, P.L., Gionanni, Y., Battaglia, F.P., Wiener, S.I., 2010. Coherent theta oscillations and reorganization of spike timing in the hippocampal-prefrontal network upon learning. *Neuron* 66, 921–936. doi:10.1016/j.neuron.2010.05.013
- Benucci, A., Frazor, R., Carandini, M., 2007. Standing waves and traveling waves distinguish two circuits in visual cortex. *Neuron* 55, 103–117.
- Berry, E., Hampshire, a, Rowe, J., Hodges, S., Kapur, N., Watson, P., Browne, G., Smyth, G., Wood, K., Owen, a M., 2009. The neural basis of effective memory therapy in a patient with limbic encephalitis. *J. Neurol. Neurosurg. Psychiatry* 80, 1202–5. doi:10.1136/jnnp.2008.164251
- Besson, G., Ceccaldi, M., Didic, M., Barbeau, E.J., 2012. The speed of visual recognition memory. *Vis. cogn.* 20, 1131–1152. doi:10.1080/13506285.2012.724034
- Besson, G., Ceccaldi, M., Tramon, E., Felician, O., Didic, M., Barbeau, E.J., 2015. Fast, but not slow, familiarity is preserved in patients with amnesic mild cognitive impairment. *Cortex* 65, 36–49. doi:10.1016/j.cortex.2014.10.020
- Bi, G., Poo, M., 1998. Synaptic modifications in cultured hippocampal neurons: dependence on spike timing, synaptic strength, and postsynaptic cell type. *J. Neurosci.* 18, 10464–10472.
- Bichler, O., Querlioz, D., Thorpe, S.J., Bourgoin, J.-P., Gamrat, C., 2012. Extraction of temporally correlated features from dynamic vision sensors with spike-timing-dependent plasticity. *Neural Netw.* 32, 339–48. doi:10.1016/j.neunet.2012.02.022
- Bichler, O., Querlioz, D., Thorpe, S.J., Bourgoin, J.-P., Gamrat, C., 2011. Unsupervised features extraction from asynchronous silicon retina through Spike-Timing-Dependent Plasticity. 2011 Int. Jt. Conf. Neural Networks 859–866. doi:10.1109/IJCNN.2011.6033311
- Binder, J.R., Desai, R.H., Graves, W.W., Conant, L.L., 2009. Where is the semantic system? A critical review and meta-analysis of 120 functional neuroimaging studies. *Cereb. Cortex* 19, 2767–96. doi:10.1093/cercor/bhp055
- Binzegger, T., Douglas, R.J., Martin, K.A.C., 2004. A quantitative map of the circuit of cat primary visual cortex. *J. Neurosci.* 24, 8441–53. doi:10.1523/JNEUROSCI.1400-04.2004
- Blumenfeld, R., Parks, C., 2011. Putting the pieces together: the role of dorsolateral prefrontal cortex in relational memory encoding. *J. Cogn. Neurosci.* 23, 257–265. doi:10.1162/jocn.2010.21459
- Bohm, D., 1989. Meaning and Information, in: Pylykänen, P. (Ed.), *The Search for Meaning* - “The New Spirit in Science and Philosophy. Crucible, pp. 1–24.
- Boldini, A., Russo, R., Avons, S.E., 2004. One process is not enough! A speed-accuracy tradeoff study of recognition memory. *Psychon. Bull. Rev.* 11, 353–361. doi:10.3758/BF03196582

- Bower, G., Forgas, J., 2000. Affect , memory , and social cognition, in: Eich, E., Kihlstrom, J., Bower, G.H., Forgas, J., Niedenthal, P. (Eds.), *Cognition and Emotion*. Oxford University Press.
- Bowers, J.S., 2009. On the biological plausibility of grandmother cells: implications for neural network theories in psychology and neuroscience. *Psychol. Rev.* 116, 220–51. doi:10.1037/a0014462
- Bowers, J.S., 2002. Challenging the widespread assumption that connectionism and distributed representations go hand-in-hand. *Cogn. Psychol.* 45, 413–445. doi:10.1016/S0010-0285(02)00506-6
- Braak, H., Del Tredici, K., 2012. Alzheimer’s disease: pathogenesis and prevention. *Alzheimers. Dement.* 8, 227–33. doi:10.1016/j.jalz.2012.01.011
- Brady, T.F., Konkle, T., Alvarez, G.A., Oliva, A., 2008. Visual long-term memory has a massive storage capacity for object details. *Proc. Natl. Acad. Sci. U. S. A.* 105, 14325–9. doi:10.1073/pnas.0803390105
- Brady, T.F., Konkle, T., Gill, J., Oliva, A., Alvarez, G.A., 2013. Visual long-term memory has the same limit on fidelity as visual working memory. *Psychol. Sci.* 24, 981–90. doi:10.1177/0956797612465439
- Brandt, K.R., Eysenck, M.W., Nielsen, M.K., von Oertzen, T.J., 2016. Selective lesion to the entorhinal cortex leads to an impairment in familiarity but not recollection. *Brain Cogn.* 104, 82–92. doi:10.1016/j.bandc.2016.02.005
- Brette, R., 2015. Philosophy of the Spike: Rate-Based vs. Spike-Based Theories of the Brain. *Front. Syst. Neurosci.* 9, 151. doi:10.3389/fnsys.2015.00151
- Buchanan, T.W., 2007. Retrieval of emotional memories. *Psychol. Bull.* 133, 761–779.
- Burgess, N., Becker, S., King, J.A., O’Keefe, J., 2001. Memory for events and their spatial context: models and experiments. *Philos. Trans. R. Soc. Lond. B. Biol. Sci.* 356, 1493–503. doi:10.1098/rstb.2001.0948
- Burianova, H., McIntosh, A.R., Grady, C.L., 2010. A common functional brain network for autobiographical, episodic, and semantic memory retrieval. *Neuroimage* 49, 865–74. doi:10.1016/j.neuroimage.2009.08.066
- Butter, C.M., Trobe, J.D., Foster, N.L., Berent, S., 1996. Visual-Spatial Deficits Explain Visual Symptoms in Alzheimer’s Disease. *Am. J. Ophthalmol.* 122, 97–105. doi:10.1016/S0002-9394(14)71969-5
- Buzsáki, G., 1989. Two-stage model of memory trace formation: a role for noisy brain states. *Neuroscience* 31, 551–570.
- Buzsáki, G., Leung, L., Vanderwolf, C., 1983. Cellular bases of hippocampal EEG in the behaving rat. *Brain Res* 287, 139–71.

- Cabeza, R., Ciaramelli, E., Olson, I.R., Moscovitch, M., 2008. The parietal cortex and episodic memory: an attentional account. *Nat. Rev. Neurosci.* 9, 613–625. doi:10.1038/nrn2459
- Capurro, A., Quiroga, R.Q., 2009. The neural correlates of perceptual awareness. *Psyche (Stuttg)*. 15, 29–38.
- Carlson, T., Tovar, D.A., Kriegeskorte, N., 2013. Representational dynamics of object vision : The first 1000 ms 13, 1–19. doi:10.1167/13.10.1.doi
- Carone, B.R., Fauquier, L., Habib, N., Shea, J.M., Caroline, E., Li, R., Bock, C., Li, C., Zamore, P.D., Weng, Z., Hofmann, H.A., Friedman, N., Rando, O.J., 2010. Paternally-induced transgenerational environmental reprogramming of metabolic gene expression in mammals. *Cell* 143, 1084–1096. doi:10.1016/j.cell.2010.12.008.Paternally-induced
- Carpenter, A.C., Schacter, D.L., 2016. Flexible retrieval: When true inferences produce false memories. *J. Exp. Psychol. Learn. Mem. Cogn.* doi:10.1037/xlm0000340
- Cauchoix, M., Arslan, A.B., Fize, D., Serre, T., 2012. The Neural Dynamics of Visual Processing in Monkey Extrastriate Cortex : A Comparison between Univariate and Multivariate Techniques 164–171.
- Cauchoix, M., Barragan-Jason, G., Serre, T., Barbeau, E.J., 2014. The neural dynamics of face detection in the wild revealed by MVPA. *J. Neurosci.* 34, 846–854. doi:10.1523/JNEUROSCI.3030-13.2014
- Cauchoix, M., Crouzet, S.M., Fize, D., Serre, T., 2016. Fast ventral stream neural activity enables rapid visual categorization. *Neuroimage* 125, 280–290. doi:10.1016/j.neuroimage.2015.10.012
- Cepeda, N.J., Pashler, H., Vul, E., Wixted, J.T., Rohrer, D., 2006. Distributed practice in verbal recall tasks: A review and quantitative synthesis. *Psychol. Bull.* 132, 354–380. doi:10.1037/0033-2909.132.3.354
- Charles, L., Van Opstal, F., Marti, S., Dehaene, S., 2013. Distinct brain mechanisms for conscious versus subliminal error detection. *Neuroimage* 73, 80–94. doi:10.1016/j.neuroimage.2013.01.054
- Chladni, E., 1787. *Entdeckungen über die Theorie des Klages.*
- Cichy, R.M., Pantazis, D., Oliva, A., 2014. Resolving human object recognition in space and time. *Nat. Neurosci.* 17, 455–462. doi:10.1038/nn.3635
- Cichy, R.M., Ramirez, F.M., Pantazis, D., 2015. Can visual information encoded in cortical columns be decoded from magnetoencephalography data in humans? *Neuroimage* 121, 193–204. doi:10.1016/j.neuroimage.2015.07.011
- Clark, L.R., Stricker, N.H., Libon, D.J., Delano-Wood, L., Salmon, D.P., Delis, D.C., Bondi, M.W., 2012. Yes/No versus forced-choice recognition memory in mild cognitive impairment and Alzheimer's disease: Patterns of impairment and associations with dementia severity. *Clin Neuropsychol* 26, 1201–1216. doi:10.1080/13854046.2012.728626.

- Cohen, M., 2014. Analyzing neural time series data: theory and practice. MIT Press.
- Conway, M.A., 2009. Episodic memories. *Neuropsychologia* 47, 2305–13. doi:10.1016/j.neuropsychologia.2009.02.003
- Conway, M.A., Cohen, G., Stanhope, N., 1991. On the Very Long-Term Retention of Knowledge Acquired Through Formal Education: Twelve Years of Cognitive Psychology. *J. Exp. Psychol. Gen.* 120, 395–409.
- Conway, M.A., Gardiner, J.M., Perfect, T.J., Anderson, S.J., Cohen, G.M., 1997. Changes in memory awareness during learning: The acquisition of knowledge by psychology undergraduates. *J. Exp. Psychol. Gen.* 126, 393–413. doi:10.1037//0096-3445.126.4.393
- Conway, M.A., Loveday, C., 2015. Remembering, imagining, false memories & personal meanings. *Conscious. Cogn.* 33, 574–81. doi:10.1016/j.concog.2014.12.002
- Conway, M.A., Pleydell-Pearce, C.W., 2000. The construction of autobiographical memories in the self-memory system. *Psychol. Rev.* 107, 261–288. doi:10.1037//0033-295X.107.2.261
- Craik, F., 1986. A functional account of age differences in memory, in: F. Klix and H. Hagendorf (Ed.), *Human Memory and Cognitive Capabilities: Mechanisms and Performances*. Elsevier Science Publishers, pp. 409–421.
- Crick, F., Koch, C., 1990. Towards a neurobiological theory of consciousness. *Semin. Neurosci.* 2, 263–275.
- Crick, F.C., 1984. Memory and Molecular turnover. *Nature* 312, 101.
- Cromer, J.A., Roy, J.E., Miller, E.K., 2010. Representation of Multiple , Independent Categories in the Primate Prefrontal Cortex. *Neuron* 66, 796–807. doi:10.1016/j.neuron.2010.05.005
- Crouzet, S.M., Busch, N.A., Ohla, K., 2015. Taste Quality Decoding Parallels Taste Sensations. *Curr. Biol.* 890–896. doi:10.1016/j.cub.2015.01.057
- Cumming, G., 2014. The new statistics: why and how. *Psychol. Sci.* 25, 7–29. doi:10.1177/0956797613504966
- Curran, E., 2003. Learning to control brain activity: A review of the production and control of EEG components for driving brain–computer interface (BCI) systems. *Brain Cogn.* 51, 326–336. doi:10.1016/S0278-2626(03)00036-8
- Curran, T., 2000. Brain potentials of recollection and familiarity. *Mem. Cognit.* 28, 923–38.
- Curran, T., Friedman, W.J., 2004. ERP old/new effects at different retention intervals in recency discrimination tasks. *Cogn. Brain Res.* 18, 107–120. doi:10.1016/j.cogbrainres.2003.09.006
- D’amour, J., Froemke, R., 2015. Inhibitory and excitatory spike-timing-dependent plasticity in the auditory cortex. *Neuron* 86, 514–528. doi:10.1016/j.neuron.2015.03.014.

- Daoudal, G., Debanne, D., 2003. Long-term plasticity of intrinsic excitability: learning rules and mechanisms. *Learn. Mem.* 10, 456–465. doi:10.1101/lm.64103
- Davis, M.H., Gaskell, M.G., 2009. A complementary systems account of word learning: neural and behavioural evidence. *Philos. Trans. R. Soc. Lond. B. Biol. Sci.* 364, 3773–800. doi:10.1098/rstb.2009.0111
- De Baene, W., Vogels, R., 2010. Effects of adaptation on the stimulus selectivity of macaque inferior temporal spiking activity and local field potentials. *Cereb. cortex* 20, 2145–65. doi:10.1093/cercor/bhp277
- De Chastelaine, M., Friedman, D., Cycowicz, Y.M., Horton, C., 2009. Effects of multiple study–test repetition on the neural correlates of recognition memory: ERPs dissociate remembering and knowing. *Psychophysiology* 46, 86–99. doi:10.1111/j.1469-8986.2008.00754.x.Effects
- De Lucia, M., Clarke, S., Murray, M.M., 2010. A temporal hierarchy for conspecific vocalization discrimination in humans. *J. Neurosci.* 30, 11210–21. doi:10.1523/JNEUROSCI.2239-10.2010
- De Meo, R., Bourquin, N.M.-P., Knebel, J.-F., Murray, M.M., Clarke, S., 2015. From bird to sparrow: Learning-induced modulations in fine-grained semantic discrimination. *Neuroimage* 118, 163–73. doi:10.1016/j.neuroimage.2015.05.091
- De Vanssay-Maigne, A., Noulhiane, M., Devauchelle, A.D., Rodrigo, S., Baudoin-Chial, S., Meder, J.F., Oppenheim, C., Chiron, C., Chassoux, F., 2011. Modulation of encoding and retrieval by recollection and familiarity: mapping the medial temporal lobe networks. *Neuroimage* 58, 1131–8. doi:10.1016/j.neuroimage.2011.06.086
- Dehaene, S., 2009. Conscious and nonconscious processes: distinct forms of evidence accumulation. *Séminaire Poincaré* XII 89–114.
- Dehaene, S., 2008. Symbols and quantities in parietal cortex : elements of a mathematical theory of number representation and manipulation, in: *Attention and Performance*. pp. 527–574.
- Dehaene, S., Kerszberg, M., Changeux, J.-P., 1998a. A neuronal model of a global workspace in effortful cognitive tasks. *Proc. Natl. Acad. Sci.* 95, 14529–14534.
- Dehaene, S., Naccache, L., Le Clec'H, G., Koechlin, E., Mueller, M., Dehaene-Lambertz, G., van de Moortele, P.-F., Le Bihan, D., 1998b. Imaging unconscious semantic priming. *Nature* 395, 597–600.
- Del Cul, A., Baillet, S., Dehaene, S., 2007. Brain dynamics underlying the nonlinear threshold for access to consciousness. *PLoS Biol.* 5, e260. doi:10.1371/journal.pbio.0050260
- Delorme, A., Makeig, S., 2004. EEGLAB: an open source toolbox for analysis of single-trial EEG dynamics including independent component analysis. *J. Neurosci. Methods* 134, 9–21. doi:10.1016/j.jneumeth.2003.10.009
- Desimone, R., 1996. Neural mechanisms for visual memory and their role in attention. *Proc. Natl. Acad. Sci. USA* 93, 13494–13499.

- Dewhurst, S.A., Anderson, S.J., 1999. Effects of exact and category repetition in true and false recognition memory. *Mem. Cognit.* 27, 665–673. doi:10.3758/BF03211560
- Dias, B.G., Ressler, K.J., 2014. Parental olfactory experience influences behavior and neural structure in subsequent generations. *Nat. Neurosci.* 17, 89–96.
- DiCarlo, J.J., Cox, D.D., 2007. Untangling invariant object recognition. *Trends Cogn. Sci.* 11, 333–41. doi:10.1016/j.tics.2007.06.010
- Didic, M., Barbeau, E.J., Felician, O., Tramon, E., Guedj, E., Poncet, M., Ceccaldi, M., 2011. Which memory system is impaired first in Alzheimer's disease? *J. Alzheimers. Dis.* 27, 11–22. doi:10.3233/JAD-2011-110557
- Didic, M., Felician, O., Barbeau, E.J., Mancini, J., Latger-Florence, C., Tramon, E., Ceccaldi, M., 2013. Impaired visual recognition memory predicts Alzheimer's disease in amnesic mild cognitive impairment. *Dement. Geriatr. Cogn. Disord.* 35, 291–9. doi:10.1159/000347203
- Diekelmann, S., Born, J., 2010. The memory function of sleep. *Nat. Rev. Neurosci.* 11, 114–26. doi:10.1038/nrn2762
- Doherty, A., Pauly-Takacs, K., Caprani, N., Gurrin, C., Moulin, C.J.A., O'Connor, N., Smeaton, A., 2012. Experiences of aiding autobiographical memory using the SenseCam. *Human-Computer Interact.* 27, 151–174.
- Donders, F.C., 1969. On the speed of mental processes. *Acta Psychol. (Amst).* 30, 412–431.
- Doppelmayr, M., Klimesch, W., Schwaiger, J., Auinger, P., Winkler, T., 1998. Theta synchronization in the human EEG and episodic retrieval. *Neurosci. Lett.* 257, 41–44. doi:10.1016/S0304-3940(98)00805-2
- Duarte, A., Ranganath, C., Winward, L., Hayward, D., Knight, R.T., 2004. Dissociable neural correlates for familiarity and recollection during the encoding and retrieval of pictures. *Cogn. Brain Res.* 18, 255–272. doi:10.1016/j.cogbrainres.2003.10.010
- Dudai, Y., 2012. The Restless Engram: Consolidations Never End. *Annu. Rev. Neurosci.* 35, 227–247. doi:10.1146/annurev-neuro-062111-150500
- Dudai, Y., 2006. Reconsolidation: the advantage of being refocused. *Curr. Opin. Neurobiol.* 16, 174–8. doi:10.1016/j.conb.2006.03.010
- Dudai, Y., 2004. The neurobiology of consolidations, or, how stable is the engram? *Annu. Rev. Psychol.* 55, 51–86. doi:10.1146/annurev.psych.55.090902.142050
- Dudai, Y., Eisenberg, M., 2004. Rites of Passage of the Engram : Reconsolidation and the Lingering Consolidation Hypothesis. *Neuron* 44, 93–100.
- Duncan, J., 2001. An adaptive coding model of neural function in prefrontal cortex. *Nat. Rev. Neurosci.* 2, 820–829.

- Duncan, K.K., Hadjipapas, A., Li, S., Kourtzi, Z., Bagshaw, A., Barnes, G., 2010. Identifying spatially overlapping local cortical networks with MEG. *Hum. Brain Mapp.* 31, 1003–16. doi:10.1002/hbm.20912
- Dunn, J.C., 2004. Remember-know: a matter of confidence. *Psychol. Rev.* 111, 524–42. doi:10.1037/0033-295X.111.2.524
- Düzel, E., Habib, R., Schott, B., Schoenfeld, a., Lobaugh, N., McIntosh, a. R., Scholz, M., Heinze, H.J., 2003. A multivariate, spatiotemporal analysis of electromagnetic time-frequency data of recognition memory. *Neuroimage* 18, 185–197. doi:10.1016/S1053-8119(02)00031-9
- Düzel, E., Yonelinas, A., Mangun, G., Heinze, H., Tulving, E., 1997. Event-related brain potential correlates of two states of conscious awareness in memory. *Proc. Natl. Acad. Sci.* 94, 5973–5978.
- Ebbinghaus, H., 1885. *Memory: A Contribution to Experimental Psychology*. New York, NY, US: Teachers College Press.
- Eichenbaum, H., Otto, T., Cohen, N.J., 2010. Two functional components of the hippocampal memory system. *Behav. Brain Sci.* 17, 449–472. doi:10.1017/S0140525X00035391
- Engel, A.K., Singer, W., 2001. Temporal binding and the neural correlates of sensory awareness. *Trends Cogn. Sci.* 5, 16–25. doi:10.1016/S1364-6613(00)01568-0
- Eschenko, O., Ramadan, W., Mölle, M., 2008. Sustained increase in hippocampal sharp-wave ripple activity during slow-wave sleep after learning. *Learn. Mem.* 15, 222–228. doi:10.1101/lm.726008.15
- Euston, D., Gruber, A., McNaughton, B., 2012. The role of medial prefrontal cortex in memory and decision making. *Neuron* 76, 1057–1070. doi:10.1016/j.neuron.2012.12.002.
- Fan, R., Wang, X., Lin, C., 2008. LIBLINEAR : A Library for Large Linear Classification 9, 1871–1874.
- Ferguson, R., Homa, D., Ellis, D., 2016. Memory for temporally dynamic scenes. *Q. J. Exp. Psychol.* 0218, 1–14. doi:10.1080/17470218.2016.1174721
- Folstein, M., Folstein, S., McHugh, P.R., 1975. A practical state method for grading the cognitive state of patients for the clinician. *J. Psychiat. Res* 12, 189–198.
- Forster, K.I., Chambers, S.M., 1973. Lexical access and naming time. *J. Verbal Learning Verbal Behav.* 12, 627–635.
- Forstmann, B.U., Keuken, M., Alkemade, A., 2015. An Introduction to Human Brain Anatomy, in: Forstmann, B.U., Wagenmakers, E.-J. (Eds.), *An Introduction to Model-Based Cognitive Neuroscience*. Springer New York, pp. 71–89. doi:10.1007/978-1-4939-2236-9
- Foster, D.J., Wilson, M. a, 2006. Reverse replay of behavioural sequences in hippocampal place cells during the awake state. *Nature* 440, 680–3. doi:10.1038/nature04587

- Frankland, P.W., Bontempi, B., 2005. The organization of recent and remote memories. *Nat. Rev. Neurosci.* 6, 119–130. doi:10.1038/nrn1607
- Frankland, P.W., Köhler, S., Josselyn, S.A., 2013. Hippocampal neurogenesis and forgetting. *Trends Neurosci.* 36, 497–503. doi:10.1016/j.tins.2013.05.002
- Freed, D., Corkin, S., 1988. Rate of forgetting in HM: 6-month recognition. *Behav. Neurosci.* 102, 823–827.
- Freedman, D.J., Riesenhuber, M., Poggio, T., Miller, E.K., 2006. Experience-dependent sharpening of visual shape selectivity in inferior temporal cortex. *Cereb. cortex* 16, 1631–44. doi:10.1093/cercor/bhj100
- Friedman, D., Chastelaine, M. de, Nessler, D., Malcolm, B., 2010. Changes in familiarity and recollection across the lifespan: an ERP perspective. *Brain Res.* 1310, 124–141. doi:10.1016/j.brainres.2009.11.016.Changes
- Friston, K., 2005. A theory of cortical responses. *Philos. Trans. R. Soc. Lond. B. Biol. Sci.* 360, 815–36. doi:10.1098/rstb.2005.1622
- Fuentemilla, L., Penny, W.D., Cashdollar, N., Bunzeck, N., Düzel, E., 2010. Theta-coupled periodic replay in working memory. *Curr. Biol.* 20, 606–12. doi:10.1016/j.cub.2010.01.057
- Funahashi, S., Bruce, C.J., Goldman-Rakic, P.S., 1989. Mnemonic coding of visual space in the monkey's dorsolateral prefrontal cortex. *J. Neurophysiol.* 61, 331–349.
- Furman, O., Dorfman, N., Hasson, U., Davachi, L., Dudai, Y., 2007. They saw a movie: long-term memory for an extended audiovisual narrative. *Learn. Mem.* 14, 457–67. doi:10.1101/lm.550407
- Fuster, J., Alexander, G., 1971. Neuron activity related to short-term memory. *Science* (80-.). 173, 652–654.
- Gabrieli, J.D.E., Cohen, N.J., Corkin, S., 1988. The impaired learning of semantic knowledge following bilateral medial temporal-lobe resection. *Brain Cogn.* 7, 157–177. doi:10.1016/0278-2626(88)90027-9
- Gais, S., Plihal, W., Wagner, U., Born, J., 2000. Early sleep triggers memory for early visual discrimination skills. *Nat. Neurosci.* 3, 1335–1339.
- Gardiner, J.M., Kaminska, Z., Dixon, M., Java, R.I., 1996. Repetition of previously novel melodies sometimes increases both remember and know responses in recognition memory. *Psychon. Bull. Rev.* 3, 366–71. doi:10.3758/BF03210762
- Gardiner, J.M., Richardson-Klavehn, A., 2000. Remembering and knowing., in: Tulving, E., Craik, F.I.M. (Eds.), *The Oxford Handbook of Memory*. pp. 229–244.
- Gardner, E., 1987. Maximum Storage Capacity in Neural Networks. *Europhys. Lett.* 4, 481–485.

- Gautrais, J., Thorpe, S., 1998. Rate coding versus temporal order coding: a theoretical approach. *Biosystems* 48, 57–65. doi:10.1016/S0303-2647(98)00050-1
- Geisler, W.S., Kersten, D., 2002. Illusions, perception and Bayes. *Nat. Neurosci.* 5, 508–10. doi:10.1038/nn0602-508
- Gelbard-Sagiv, H., Mukamel, R., Harel, M., 2008. Internally generated reactivation of single neurons in human hippocampus during free recall. *Science* (80-.). 322, 96–101. doi:10.1126/science.1164685.Internally
- Genon, S., Simon, J., Bahri, M.A., Collette, F., Souchay, C., Jaspar, M., Bastin, C., Salmon, E., 2016. Relating pessimistic memory predictions to Alzheimer’s disease brain structure. *Cortex* 85, 151–164. doi:10.1016/j.cortex.2016.09.014
- Girardeau, G., Benchenane, K., Wiener, S.I., Buzsáki, G., Zugaro, M.B., 2009. Selective suppression of hippocampal ripples impairs spatial memory. *Nat. Neurosci.* 12, 1222–1223. doi:10.1038/nn.2384
- Gisquet-Verrier, P., Riccio, D.C., 2012. Memory reactivation effects independent of reconsolidation. *Learn. Mem.* 19, 401–9. doi:10.1101/lm.026054.112
- Gold, J.I., Shadlen, M.N., 2007. The neural basis of decision making. *Annu. Rev. Neurosci.* 30, 535–74. doi:10.1146/annurev.neuro.29.051605.113038
- Goldman-Rakic, P.S., 1987. Circuitry of Primate Prefrontal Cortex and Regulation of Behavior by Representational Memory, in: *Handbook of Physiology*, Vol 5. pp. 373–417.
- Gomez-Isla, T., Price, J.L., Mckeel Jr., D.W., Morris, J.C., Growdon, J.H., Hyman, B.T., 1996. Profound Loss of Layer II Entorhinal Cortex Neurons Occurs in Very Mild Alzheimer’s Disease. *J. Neurosci.* 16, 4491–4500.
- Graham, D., Field, D., 2006. Sparse coding in the neocortex, in: Kaas, J.H. (Ed.), *Evolution of Nervous Systems*. Oxford: Academic Press, pp. 181–187.
- Grill-Spector, K., Henson, R., Martin, A., 2006. Repetition and the brain: neural models of stimulus-specific effects. *Trends Cogn. Sci.* 10, 14–23. doi:10.1016/j.tics.2005.11.006
- Gross, C.G., 2002. Genealogy of the “Grandmother Cell”. *Neurosci.* 8, 512–518. doi:10.1177/107385802237175
- Grossman, A., Morlet, J., 1984. Decomposition of Hardy functions into square integrable wavelets of constant shape. *SIAM J. Math. Anal.* 15, 723–736.
- Gruber, T., Tsivilis, D., 2008. Induced electroencephalogram oscillations during source memory: familiarity is reflected in the gamma band, recollection in the theta band. *J. Cogn. Neurosci.* 20, 1043–1053.
- Habib, R., Nyberg, L., 2008. Neural correlates of availability and accessibility in memory. *Cereb. Cortex* 18, 1720–6. doi:10.1093/cercor/bhm201

- Halgren, E., Marinkovic, K., Chauvel, P., 1998. Generators of the late cognitive potentials in auditory and visual oddball tasks. *Electroencephalogr. Clin. Neurophysiol.* 106, 156–164. doi:10.1016/S0013-4694(97)00119-3
- Hanslmayr, S., Staudigl, T., 2014. How brain oscillations form memories--a processing based perspective on oscillatory subsequent memory effects. *Neuroimage* 85 Pt 2, 648–55. doi:10.1016/j.neuroimage.2013.05.121
- Hanslmayr, S., Staudigl, T., Fellner, M.-C., 2012. Oscillatory power decreases and long-term memory: the information via desynchronization hypothesis. *Front. Hum. Neurosci.* 6, 74. doi:10.3389/fnhum.2012.00074
- Harand, C., Bertran, F., La Joie, R., Landeau, B., Mézenge, F., Desgranges, B., Peigneux, P., Eustache, F., Rauchs, G., 2012. The hippocampus remains activated over the long term for the retrieval of truly episodic memories. *PLoS One* 7, e43495. doi:10.1371/journal.pone.0043495
- Hardt, O., Nader, K., Nadel, L., 2013. Decay happens: the role of active forgetting in memory. *Trends Cogn. Sci.* 17, 111–20. doi:10.1016/j.tics.2013.01.001
- Hart, J.T., 1965. Memory and the feeling-of-knowing experience. *J. Educ. Psychol.* 56, 208–216.
- Hasselmo, M.E., Stern, C.E., 2014. Theta rhythm and the encoding and retrieval of space and time. *Neuroimage* 85, 656–66. doi:10.1016/j.neuroimage.2013.06.022
- Haxby, J., 2012. Multivariate pattern analysis of fMRI: the early beginnings. *Neuroimage* 62, 852–855. doi:10.1016/j.neuroimage.2012.03.016.Multivariate
- Haxby, J., Gobbini, M., Furey, M., Schouten, J., Pietrini, P., 2001. Distributed and overlapping representations of faces and objects in ventral temporal cortex. *Science* (80-.). 293, 2425–2430.
- Hebb, D.O., 1949. *The organization of behavior. A neuropsychological theory.* New York.
- Helmholtz, H. von, 1947. Concerning the perception in general, 1867, in: Wayne, D. (Ed.), *Readings in the History of Psychology.* Appleton-Century-Croft, pp. 214–230.
- Herbert, D.M.B., Burt, J.S., 2004. What do students remember? Episodic memory and the development of schematization. *Appl. Cogn. Psychol.* 18, 77–88. doi:10.1002/acp.947
- Hodges, S., Williams, L., Berry, E., 2006. SenseCam: A retrospective memory aid. *Ubicomp'06* 177–193.
- Hopfield, J.J., 1982. Neural networks and physical systems with emergent collective computational abilities. *Proc. Natl. Acad. Sci.* 79, 2554–2558. doi:10.1073/pnas.79.8.2554
- Iaccarino, H.F., Singer, A.C., Martorell, A.J., Rudenko, A., Gao, F., Gillingham, T.Z., Mathys, H., Seo, J., Kritskiy, O., Abdurrob, F., Adaikkan, C., Canter, R.G., Rueda, R., Brown, E.N., Boyden, E.S., Tsai, L.-H., 2016. Gamma frequency entrainment attenuates amyloid load and modifies microglia. *Nature* 540, 230–235. doi:10.1038/nature20587

- Isik, L., Meyers, E.M., Leibo, J.Z., Poggio, T., 2014. The dynamics of invariant object recognition in the human visual system. *J. Neurophysiol.* 111, 91–102. doi:10.1152/jn.00394.2013
- Ison, M.J., Mormann, F., Cerf, M., Koch, C., Fried, I., Quiroga, R.Q., 2011. Selectivity of pyramidal cells and interneurons in the human medial temporal lobe. *J. Neurophysiol.* 106, 1713–21. doi:10.1152/jn.00576.2010
- Ison, M.J., Quiroga, R., Fried, I., 2015. Rapid Encoding of New Memories by Individual Neurons in the Human Brain. *Neuron* 87, 220–30. doi:10.1016/j.neuron.2015.06.016
- Jacobs, J., 2014. Hippocampal theta oscillations are slower in humans than in rodents: implications for models of spatial navigation and memory. *Phil. Trans. R. Soc. B* 369, 20130304.
- Jacobs, J., Hwang, G., Curran, T., Kahana, M.J., 2006. EEG oscillations and recognition memory: theta correlates of memory retrieval and decision making. *Neuroimage* 32, 978–87. doi:10.1016/j.neuroimage.2006.02.018
- Jacoby, L., 1991. A process dissociation framework: Separating automatic from intentional uses of memory. *J. Mem. Lang.* 30, 513–541.
- Jacoby, L., Whitehouse, K., 1989. An illusion of memory: False recognition influenced by unconscious perception. *J. Exp. Psychol. Gen.* 118, 126–135. doi:10.1037//0096-3445.118.2.126
- Jacoby, L.L., Dallas, M., 1981. On the relationship between autobiographical memory and perceptual learning. *J. Exp. Psychol. Gen.* 110, 306–340. doi:10.1037//0096-3445.110.3.306
- Jadhav, S.P., Kemere, C., German, P.W., Frank, L.M., 2012. Awake hippocampal sharp-wave ripples support spatial memory. *Science* (80-.). 336, 1454–1458. doi:10.1126/science.1217230.Awake
- Jaeger, A., Johnson, J.D., Corona, M., Rugg, M.D., 2010. ERP correlates of the incidental retrieval of emotional information: effects of study-test delay. *Brain Res.* 1269, 105–113. doi:10.1016/j.brainres.2009.02.082.ERP
- Jafarpour, A., Fuentemilla, L., Horner, A.J., Penny, W., Duzel, E., 2014. Replay of very early encoding representations during recollection. *J. Neurosci.* 34, 242–8. doi:10.1523/JNEUROSCI.1865-13.2014
- Jafarpour, A., Horner, A.J., Fuentemilla, L., Penny, W.D., Duzel, E., 2013. Decoding oscillatory representations and mechanisms in memory. *Neuropsychologia* 51, 772–80. doi:10.1016/j.neuropsychologia.2012.04.002
- Jiang, H., Gerven, M. van, Jensen, O., 2015. Modality-specific alpha modulations facilitate long-term memory encoding in the presence of distracters. *J. Cogn. Neurosci.* 27, 583–592. doi:10.1162/jocn
- Johnson, M., Hashtroudi, S., Lindsay, D., 1993. Source monitoring. *Psychol. Bull.* 114, 3–28.

- Joubert, O.R., Rousselet, G.A., Fize, D., Fabre-Thorpe, M., 2007. Processing scene context: fast categorization and object interference. *Vision Res.* 47, 3286–97. doi:10.1016/j.visres.2007.09.013
- Jung, H., Francis-west, P.H., Widelitz, R.B., Jiang, T., Ting-berreth, S., Tickle, C., Wolpert, L., Chuong, C., 1998. Local Inhibitory Action of BMPs and Their Relationships with Activators in Feather Formation: Implications for Periodic Patterning. *Dev. Biol.* 196, 11–23.
- Kamitani, Y., Tong, F., 2005. Decoding the visual and subjective contents of the human brain. *Nat. Neurosci.* 8, 679–85. doi:10.1038/nn1444
- Karlsson, M., Frank, L., 2009. Awake replay of remote experiences in the hippocampus. *Nat. Neurosci.* 12, 913–918. doi:10.1038/nn.2344.Awake
- Kersten, D., Mamassian, P., Yuille, A., 2004. Object perception as Bayesian inference. *Annu. Rev. Psychol.* 55, 271–304. doi:10.1146/annurev.psych.55.090902.142005
- Khan, U.A., Liu, L., Provenzano, F.A., Berman, D.E., Profaci, C.P., Sloan, R., Mayeux, R., Duff, K.E., Small, S.A., 2014. Molecular drivers and cortical spread of lateral entorhinal cortex dysfunction in preclinical Alzheimer's disease. *Nat. Neurosci.* 17, 304–311. doi:10.1038/nn.3606.Molecular
- Kim, Y.J., Park, S.W., Yeom, H.G., Bang, M.S., Kim, J.S., Chung, C.K., Kim, S., 2015. A study on a robot arm driven by three-dimensional trajectories predicted from non-invasive neural signals. *Biomed. Eng. Online* 14, 81. doi:10.1186/s12938-015-0075-8
- King, J.-R., Dehaene, S., 2014. A model of subjective report and objective discrimination as categorical decisions in a vast representational space. *Philos. Trans. R. Soc. Lond. B. Biol. Sci.* 369, 20130204.
- King, J.-R., Dehaene, S., 2014. Characterizing the dynamics of mental representations: the temporal generalization method. *Trends Cogn. Sci.* 18, 203–10. doi:10.1016/j.tics.2014.01.002
- King, J.-R., Gramfort, A., Schurger, A., Naccache, L., Dehaene, S., 2014. Two distinct dynamic modes subtend the detection of unexpected sounds. *PLoS One* 9, e85791. doi:10.1371/journal.pone.0085791
- Kirchner, H., Thorpe, S.J., 2006. Ultra-rapid object detection with saccadic eye movements: visual processing speed revisited. *Vision Res.* 46, 1762–76. doi:10.1016/j.visres.2005.10.002
- Klampfl, S., Maass, W., 2013. Emergence of dynamic memory traces in cortical microcircuit models through STDP. *J. Neurosci.* 33, 11515–29. doi:10.1523/JNEUROSCI.5044-12.2013
- Klimesch, W., Doppelmayr, M., Schwaiger, J., Winkler, T., Gruber, W., 2000. Theta oscillations and the ERP old/new effect: independent phenomena? *Clin. Neurophysiol.* 111, 781–793. doi:10.1016/S1388-2457(00)00254-6

- Klimesch, W., Schimke, H., Schwaiger, J., 1994. Episodic and semantic memory: an analysis in the EEG theta and alpha band. *Electroencephalogr. Clin. Neurophysiol.* 91, 428–441. doi:10.1016/0013-4694(94)90164-3
- Knill, D.C., Pouget, A., 2004. The Bayesian brain: the role of uncertainty in neural coding and computation. *Trends Neurosci.* 27, 712–9. doi:10.1016/j.tins.2004.10.007
- Ko, Y., Lau, H., 2012. A detection theoretic explanation of blindsight suggests a link between conscious perception and metacognition. *Philos. Trans. R. Soc. Lond. B. Biol. Sci.* 367, 1401–11. doi:10.1098/rstb.2011.0380
- Koriat, A., 2007. Metacognition and consciousness, in: Zelazo, P., Moscovitch, M., Thompson, E. (Eds.), *The Cambridge Handbook of Consciousness*. Cambridge University Press.
- Koriat, A., 1995. Dissociating knowing and the feeling of knowing: Further evidence for the accessibility model. *J. Exp. Psychol. Gen.* 124, 311–333. doi:10.1037/0096-3445.124.3.311
- Koriat, A., Levy-Sadot, R., 2001. The combined contributions of the cue-familiarity and accessibility heuristics to feelings of knowing. *J. Exp. Psychol. Learn. Mem. Cogn.* 27, 34–53. doi:10.1037//0278-7393.27.1.34
- Kosslyn, S., Ganis, G., Thompson, W., 2001. Neural foundations of imagery. *Nat. Rev. Neurosci.* 2, 635–642.
- Kreiman, G., Fried, I., Koch, C., 2002. Single-neuron correlates of subjective vision in the human medial temporal lobe. *Proc. Natl. Acad. Sci. U. S. A.* 99, 8378–83. doi:10.1073/pnas.072194099
- Kriegeskorte, N., Kievit, R.A., 2013. Representational geometry: integrating cognition, computation, and the brain. *Trends Cogn. Sci.* 17, 401–12. doi:10.1016/j.tics.2013.06.007
- Kriegeskorte, N., Mur, M., Bandettini, P., 2008. Representational similarity analysis - connecting the branches of systems neuroscience. *Front. Syst. Neurosci.* 2, 4. doi:10.3389/neuro.06.004.2008
- Lacy, J.W., Stark, C.E.L., 2013. The neuroscience of memory: Implications for the courtroom. *Nat. Rev. Neurosci.* 14, 649–658. doi:10.1038/nrn3563
- Lee, J.L.C., 2008. Memory reconsolidation mediates the strengthening of memories by additional learning. *Nat. Neurosci.* 11, 1264–6. doi:10.1038/nn.2205
- Lega, B., Burke, J., Jacobs, J., Kahana, M.J., 2016. Slow-Theta-to-Gamma Phase-Amplitude Coupling in Human Hippocampus Supports the Formation of New Episodic Memories. *Cereb. cortex* 26, 268–78. doi:10.1093/cercor/bhu232
- Lennie, P., 2003. The cost of cortical computation. *Curr. Biol.* 13, 493–497. doi:10.1016/S
- LePort, A., Mattfeld, A., Dickinson-Anson, H., Fallon, J., Stark, C.E.L., Kruggel, F., Cahill, L., McGaugh, J.L., 2012. Behavioral and neuroanatomical investigation of highly

- superior autobiographical memory (HSAM). *Neurobiol. Learn. Mem.* 98, 78–92. doi:10.1016/j.nlm.2012.05.002.Behavioral
- Leutgeb, J.K., Leutgeb, S., Moser, M.-B., Moser, E.I., 2007. Pattern separation in the dentate gyrus and CA3 of the hippocampus. *Science* (80-.). 315, 961–966.
- Levine, B., Svoboda, E., Hay, J.F., Winocur, G., Moscovitch, M., 2002. Aging and autobiographical memory: Dissociating episodic from semantic retrieval. *Psychol. Aging* 17, 677–689. doi:10.1037//0882-7974.17.4.677
- Lewis, D.J., 1979. Psychobiology of Active and Inactive Memory. *Psychol. Bull.* 86, 1054–1083.
- Li, L.I.N., Miller, E.K., Desimone, R., 1993. The Representation of Stimulus Familiarity Temporal Cortex in Anterior Inferior. *J. Neurophysiol.* 69, 1918–1929.
- Lindenberger, U., Mayr, U., 2014. Cognitive aging: is there a dark side to environmental support? *Trends Cogn. Sci.* 18, 7–15. doi:10.1016/j.tics.2013.10.006
- Loftus, E., 2005. Planting misinformation in the human mind: A 30-year investigation of the malleability of memory. *Learn. Mem.* 12, 361–366. doi:10.1101/lm.94705.recalled
- Loftus, E.F., 1997. Creating False Memories. *Sci. Am.* 277, 70–75. doi:10.1038/scientificamerican0997-70
- Logothetis, N.K., 1998. Single units and conscious vision. *Philos. Trans. R. Soc. Lond. B. Biol. Sci.* 353, 1801–18. doi:10.1098/rstb.1998.0333
- Loschky, L.C., Larson, A.M., 2010. The natural/man-made distinction is made before basic-level distinctions in scene gist processing. *Vis. cogn.* 18, 513–536. doi:10.1080/13506280902937606
- Loveday, C., Conway, M.A., 2011. Using SenseCam with an amnesic patient: accessing inaccessible everyday memories. *Memory* 19, 697–704. doi:10.1080/09658211.2011.610803
- Macé, M.J.-M., Joubert, O.R., Nespoulous, J.-L., Fabre-Thorpe, M., 2009. The time-course of visual categorizations: you spot the animal faster than the bird. *PLoS One* 4, e5927. doi:10.1371/journal.pone.0005927
- Mack, M.L., Palmeri, T.J., 2011. The timing of visual object categorization. *Front. Psychol.* 2, 165. doi:10.3389/fpsyg.2011.00165
- Maguire, E.A., 2001. Neuroimaging studies of autobiographical event memory. *Philos. Trans. R. Soc. Lond. B. Biol. Sci.* 356, 1441–51. doi:10.1098/rstb.2001.0944
- Maguire, E.A., Frith, C.D., 2003. Lateral asymmetry in the hippocampal response to the remoteness of autobiographical memories. *J. Neurosci.* 23, 5302–5307.
- Mandler, G., 2008. Familiarity breeds attempts: A critical review of dual-process theories of recognition. *Perspect. Psychol. Sci.* 3, 390–399.

- Mandler, G., 1980. Recognizing : The Judgment of Previous Occurrence. *Psychol. Rev.* 87, 252–271.
- Mäntylä, T., Cornoldi, C., 2002. Remembering changes: repetition effects in face recollection. *Acta Psychol. (Amst)*. 109, 95–105.
- Marcel, A.J., 1983. Conscious and Unconscious Perception : Experiments Visual Masking and Word Recognition. *Cogn. Psychol.* 15, 197–237.
- Markram, H., Lubke, J., Frotscher, M., Sakmann, B., 1997. Regulation of Synaptic Efficacy by Coincidence of Postsynaptic APs and EPSPs. *Science* (80-.). 275, 1–4. doi:10.1126/science.275.5297.213
- Marr, D., 1970. Simple memory: a theory for archicortex. *Proc. R. Soc. London. Ser. B, Biol. Sci.* 262, 23–81.
- Masquelier, T., Guyonneau, R., Thorpe, S., 2009. Competitive STDP-based spike pattern learning. *Neural Comput.* 21, 1259–76. doi:10.1162/neco.2008.06-08-804
- Masquelier, T., Thorpe, S.J., 2007. Unsupervised learning of visual features through spike timing dependent plasticity. *PLoS Comput Biol* 3, e31. doi:10.1371/journal.pcbi.0030031
- McClelland, J.L., McNaughton, B.L., O'Reilly, R.C., 1995. Why there are complementary learning systems in the hippocampus and neocortex: insights from the successes and failures of connectionist models of learning and memory. *Psychol. Rev.* 102, 419–57.
- McKenzie, S., Eichenbaum, H., 2011. Consolidation and reconsolidation: two lives of memories? *Neuron* 71, 224–233. doi:10.1016/j.neuron.2011.06.037
- McKone, E., French, B., 2001. In what sense is implicit memory “episodic”? The effect of reinstating environmental context. *Psychon. Bull. Rev.* 8, 806–811. doi:10.3758/BF03196221
- Meinhardt, H., Gierer, A., 2000. Pattern formation by local self- activation and lateral inhibition. *BioEssays* 22, 753–760. doi:10.1002/1521-1878
- Mervis, C.B., Rosch, E., 1981. Categorization of Natural Objects. *Annu. Rev. Psychol.* 32, 89–115.
- Miller, E.K., 2000. The prefrontal cortex and cognitive control. *Nat. Rev. Neurosci.* 1, 59–65.
- Misanin, J.R., Miller, R.R., Lewis, D.J., 1968. Retrograde Amnesia Produced by Electroconvulsive Shock after Reactivation of a Consolidated Memory Trace. *Science* (80-.). 160, 554–555.
- Mitchell, D.B., 2006. Nonconscious priming after 17 years: invulnerable implicit memory? *Psychol. Sci.* 17, 925–9. doi:10.1111/j.1467-9280.2006.01805.x
- Miyashita, Y., 1993. Inferior temporal cortex: where visual perception meets memory. *Annu. Rev. Neurosci.* 16, 245–263.

- Miyashita, Y., 1988. Neuronal correlate of visual associative long-term memory in the primate temporal cortex. *Nature* 335, 817–820.
- Mormann, F., Kornblith, S., Quiroga, R.Q., Kraskov, A., Cerf, M., Fried, I., Koch, C., 2008. Latency and selectivity of single neurons indicate hierarchical processing in the human medial temporal lobe. *J. Neurosci.* 28, 8865–8872. doi:10.1523/JNEUROSCI.1640-08.2008.Latency
- Moscovitch, M., 1992. Memory and Working-with-Memory: A Component Process Model Based on Modules and Central Systems. *J. Cogn. Neurosci.* 4, 257–67. doi:10.1162/jocn.1992.4.3.257
- Moscovitch, M., Nadel, L., Winocur, G., Gilboa, A., Rosenbaum, R.S., 2006. The cognitive neuroscience of remote episodic, semantic and spatial memory. *Curr. Opin. Neurobiol.* 16, 179–190. doi:10.1016/j.conb.2006.03.013
- Moulin, C., Souchay, C., 2015. An active inference and epistemic value view of metacognition. *Cogn. Neurosci.* 6, 221–2. doi:10.1080/17588928.2015.1051015
- Moulin, C.J.A., 2013. Disordered recognition memory: recollective confabulation. *Cortex* 49, 1541–52. doi:10.1016/j.cortex.2013.01.010
- Moulin, C.J.A., Conway, M.A., Thompson, R.G., James, N., Jones, R.W., 2005. Disordered memory awareness: recollective confabulation in two cases of persistent déjà vecu. *Neuropsychologia* 43, 1362–78. doi:10.1016/j.neuropsychologia.2004.12.008
- Moulin, C.J.A., Souchay, C., Morris, R.G., 2013. The cognitive neuropsychology of recollection. *Cortex* 49, 1445–1451.
- Murphy, G.L., Smith, E.E., 1982. Basic-level superiority in picture categorization. *J. Verbal Learning Verbal Behav.* 21, 1–20. doi:10.1016/S0022-5371(82)90412-1
- Murray, M.M., Camen, C., Gonzalez Andino, S.L., Bovet, P., Clarke, S., 2006. Rapid brain discrimination of sounds of objects. *J. Neurosci.* 26, 1293–1302. doi:10.1523/JNEUROSCI.4511-05.2006
- Nadel, L., Moscovitch, M., 1997. Memory consolidation, retrograde amnesia and the hippocampal complex. *Curr. Opin. Neurobiol.* 7, 217–27.
- Nader, K., Schafe, G., Doux, J. Le, 2000. Fear memories require protein synthesis in the amygdala for reconsolidation after retrieval. *Nature* 406, 722–726.
- Nairne, J.S., 2002. The myth of the encoding-retrieval match. *Memory* 10, 389–95. doi:10.1080/09658210244000216
- Neisser, U., 1997. The ecological study of memory. *Philos. Trans. R. Soc. Lond. B. Biol. Sci.* 352, 1697–701. doi:10.1098/rstb.1997.0151
- Neisser, U., 1981. John Dean's memory: A case study. *Cognition* 9, 1–22.

- Nessler, B., Pfeiffer, M., Buesing, L., Maass, W., 2013. Bayesian computation emerges in generic cortical microcircuits through spike-timing-dependent plasticity. *PLoS Comput. Biol.* 9, e1003037. doi:10.1371/journal.pcbi.1003037
- Nessler, D., Friedman, D., Johnson, R., Bersick, M., 2007. Does repetition engender the same retrieval processes in young and older adults? *Neuroreport* 18, 1837–40. doi:10.1097/WNR.0b013e3282f16d9f
- Nessler, D., Mecklinger, A., 2003. ERP correlates of true and false recognition after different retention delays: Stimulus- and response-related processes. *Psychophysiology* 40, 146–159. doi:10.1111/1469-8986.00015
- Niikura, T., Tajima, H., Kita, Y., 2006. Neuronal cell death in Alzheimer's disease and a neuroprotective factor, humanin. *Curr. Neuropharmacol.* 4, 139–147.
- Nishimoto, S., Vu, A., Naselaris, T., Benjamini, Y., 2011. Reconstructing visual experiences from brain activity evoked by natural movies. *Curr. Biol.* 21, 1641–1646. doi:10.1016/j.cub.2011.08.031.Reconstructing
- Norman, K.A., O'Reilly, R.C., 2003. Modeling hippocampal and neocortical contributions to recognition memory: a complementary-learning-systems approach. *Psychol. Rev.* 110, 611–46. doi:10.1037/0033-295X.110.4.611
- Norris, D., Kinoshita, S., 2008. Perception as evidence accumulation and Bayesian inference: insights from masked priming. *J. Exp. Psychol. Gen.* 137, 434–55. doi:10.1037/a0012799
- Nyberg, L., Habib, R., McIntosh, a. R., Tulving, E., 2000. Reactivation of encoding-related brain activity during memory retrieval. *Proc. Natl. Acad. Sci.* 97, 11120–11124. doi:10.1073/pnas.97.20.11120
- Nyhus, E., Curran, T., 2010. Functional role of gamma and theta oscillations in episodic memory. *Neurosci. Biobehav. Rev.* 34, 1023–35. doi:10.1016/j.neubiorev.2009.12.014
- O'Connor, A.R., Lever, C., Moulin, C.J.A., 2010. Novel insights into false recollection: a model of déjà vécu. *Cogn. Neuropsychiatry* 15, 118–44. doi:10.1080/13546800903113071
- O'Connor, A.R., Moulin, C.J.A., 2010. Recognition without identification, erroneous familiarity, and déjà vu. *Curr. Psychiatry Rep.* 12, 165–73. doi:10.1007/s11920-010-0119-5
- O'Keefe, J., Dostrovsky, J., 1971. The hippocampus as a spatial map. Preliminary evidence from unit activity in the freely-moving rat. *Brain Res.* 34, 171–175.
- Oizumi, M., Albantakis, L., Tononi, G., 2014. From the phenomenology to the mechanisms of consciousness: Integrated Information Theory 3.0. *PLoS Comput. Biol.* 10, e1003588. doi:10.1371/journal.pcbi.1003588
- Olsen, R.K., Moses, S.N., Riggs, L., Ryan, J.D., 2012. The hippocampus supports multiple cognitive processes through relational binding and comparison. *Front. Hum. Neurosci.* 6, 1–13. doi:10.3389/fnhum.2012.00146

- Osipova, D., Takashima, A., Oostenveld, R., Fernández, G., Maris, E., Jensen, O., 2006. Theta and gamma oscillations predict encoding and retrieval of declarative memory. *J. Neurosci.* 26, 7523–31. doi:10.1523/JNEUROSCI.1948-06.2006
- Oudiette, D., Antony, J.W., Creery, J.D., Paller, K.A., 2013. The role of memory reactivation during wakefulness and sleep in determining which memories endure. *J. Neurosci.* 33, 6672–8. doi:10.1523/JNEUROSCI.5497-12.2013
- Paivio, A., Csapo, K., 1973. Picture superiority in free recall: Imagery or dual coding? *Cogn. Psychol.* 5, 176–206. doi:10.1016/0010-0285(73)90032-7
- Parker, E.S., Cahill, L., McGaugh, J.L., 2006. A case of unusual autobiographical remembering. *Neurocase* 12, 35–49. doi:10.1080/13554790500473680
- Pastötter, B., Bäuml, K.-H.T., 2014. Distinct slow and fast cortical theta dynamics in episodic memory retrieval. *Neuroimage* 94, 155–61. doi:10.1016/j.neuroimage.2014.03.002
- Patihis, L., Frenda, S.J., LePort, A.K.R., Petersen, N., Nichols, R.M., Stark, C.E.L., McGaugh, J.L., Loftus, E.F., 2013. False memories in highly superior autobiographical memory individuals. *Proc. Natl. Acad. Sci. U. S. A.* 110, 20947–52. doi:10.1073/pnas.1314373110
- Paynter, C.A., Reder, L.M., Kieffaber, P.D., 2009. Knowing we know before we know: ERP correlates of initial feeling-of-knowing. *Neuropsychologia* 47, 796–803. doi:10.1016/j.neuropsychologia.2008.12.009.Knowing
- Peigneux, P., Laureys, S., Fuchs, S., Collette, F., Perrin, F., Reggers, J., Phillips, C., Degueldre, C., Del Fio, G., Aerts, J., Luxen, A., Maquet, P., 2004. Are spatial memories strengthened in the human hippocampus during slow wave sleep? *Neuron* 44, 535–45. doi:10.1016/j.neuron.2004.10.007
- Peigneux, P., Orban, P., Balteau, E., Degueldre, C., Luxen, A., Laureys, S., Maquet, P., 2006. Offline persistence of memory-related cerebral activity during active wakefulness. *PLoS Biol.* 4, 647–658. doi:10.1371/journal.pbio.0040100
- Peirce, J.W., 2007. PsychoPy--Psychophysics software in Python. *J. Neurosci. Methods* 162, 8–13. doi:10.1016/j.jneumeth.2006.11.017
- Peyrache, A., Khamassi, M., Benchenane, K., Wiener, S.I., Battaglia, F.P., 2009. Replay of rule-learning related neural patterns in the prefrontal cortex during sleep. *Nat. Neurosci.* 12, 919–926. doi:10.1038/nn.2337
- Piolino, P., Desgranges, B., Benali, K., Eustache, F., 2002. Episodic and semantic remote autobiographical memory in ageing. *Memory* 10, 239–57. doi:10.1080/09658210143000353
- Piolino, P., Desgranges, B., Eustache, F., 2009. Episodic autobiographical memories over the course of time: cognitive, neuropsychological and neuroimaging findings. *Neuropsychologia* 47, 2314–29. doi:10.1016/j.neuropsychologia.2009.01.020
- Platt, M.L., Glimcher, P.W., 1999. Neural correlates of decision variables in parietal cortex. *Nature* 400, 233–238.

- Ploran, E.J., Nelson, S.M., Velanova, K., Donaldson, D.I., Petersen, S.E., Wheeler, M.E., 2007. Evidence accumulation and the moment of recognition: dissociating perceptual recognition processes using fMRI. *J. Neurosci.* 27, 11912–24. doi:10.1523/JNEUROSCI.3522-07.2007
- Poncet, M., Fabre-Thorpe, M., 2014. Stimulus duration and diversity do not reverse the advantage for superordinate-level representations: the animal is seen before the bird. *Eur. J. Neurosci.* 39, 1508–16. doi:10.1111/ejn.12513
- Possin, K.L., 2010. Visual spatial cognition in neurodegenerative disease. *Neurocase* 16, 466–487. doi:10.1080/13554791003730600.Visual
- Praß, M., Grimsen, C., König, M., Fahle, M., 2013. Ultra rapid object categorization: effects of level, animacy and context. *PLoS One* 8, e68051. doi:10.1371/journal.pone.0068051
- Preston, A.R., Eichenbaum, H., 2013. Interplay of hippocampus and prefrontal cortex in memory. *Curr. Biol.* 23, R764–73. doi:10.1016/j.cub.2013.05.041
- Quiroga, R.Q., 2012. Concept cells: the building blocks of declarative memory functions. *Nat. Rev. Neurosci.* 13, 587–97. doi:10.1038/nrn3251
- Quiroga, R.Q., Fried, I., Koch, C., 2013. Brain Cells for Grandmother. *Sci. Am.* 308, 30–35. doi:10.1038/scientificamerican0213-30
- Quiroga, R.Q., Kreiman, G., Koch, C., Fried, I., 2008a. Sparse but not “grandmother-cell” coding in the medial temporal lobe. *Trends Cogn. Sci.* 12, 87–91. doi:10.1016/j.tics.2007.12.003
- Quiroga, R.Q., Mukamel, R., Isham, E.A., Malach, R., Fried, I., 2008b. Human single-neuron responses at the threshold of conscious recognition. *Proc. Natl. Acad. Sci.* 105, 3599–3604.
- Quiroga, R.Q., Reddy, L., Kreiman, G., Koch, C., Fried, I., 2005. Invariant visual representation by single neurons in the human brain. *Nature* 435, 1102–7. doi:10.1038/nature03687
- Rainer, G., Asaad, W., Miller, E., 1998. Memory fields of neurons in the primate prefrontal cortex. *Proc. Natl. Acad. Sci.* 95, 15008–15013.
- Rajaram, S., Hamilton, M., Bolton, A., 2002. Distinguishing states of awareness from confidence during retrieval: Evidence from amnesia. *Cogn. Affect. Behav. Neurosci.* 2, 227–235. doi:10.3758/CABN.2.3.227
- Rasmussen, A.S., Berntsen, D., 2009. The possible functions of involuntary autobiographical memories. *Appl. Cogn. Psychol.* 23, 1137–1152. doi:10.1002/acp.1615
- Ratcliff, R., 1978. A theory of memory retrieval. *Psychol. Rev.* 85, 59–107.
- Ratcliff, R., McKoon, G., 2008. The Diffusion Decision Model: Theory and Data for Two-Choice Decision Tasks. *Neural Comput.* 20, 873–922.

- Ratcliff, R., Smith, P.L., Brown, S.D., McKoon, G., 2016. Diffusion Decision Model: Current Issues and History. *Trends Cogn. Sci.* 20, 260–281. doi:10.1016/j.tics.2016.01.007
- Rathbone, C.J., Moulin, C.J.A., Conway, M.A., 2008. Self-centered memories: the reminiscence bump and the self. *Mem. Cognit.* 36, 1403–14. doi:10.3758/MC.36.8.1403
- Reddy, L., Poncet, M., Self, M.W., Peters, J.C., Douw, L., van Dellen, E., Claus, S., Reijneveld, J.C., Baayen, J.C., Roelfsema, P.R., 2015. Learning of anticipatory responses in single neurons of the human medial temporal lobe. *Nat. Commun.* 6, 8556. doi:10.1038/ncomms9556
- Reddy, L., Thorpe, S.J., 2014. Concept cells through associative learning of high-level representations. *Neuron* 84, 248–51. doi:10.1016/j.neuron.2014.10.004
- Reder, L.M., 1987. Strategy Selection in Question Answering. *Cogn. Psychol.* 19, 90–138.
- Reder, L.M., Ritter, F.E., 1992. What determines initial feeling of knowing? Familiarity with question terms, not with the answer. *J. Exp. Psychol. Learn. Mem. Cogn.* 18, 435–451.
- Reed, H.J., 1931. The influence of a change of conditions upon the amount recalled. *J. Exp. Psychol.* 632–649.
- Reed, J.M., Squire, L.R., 1998. Retrograde Amnesia for Facts and Events: Findings from Four New Cases. *J. Neurosci.* 18, 3943–3954.
- Refaeilzadeh, P., Tang, L., Liu, H., 2009. “Cross-Validation”, in: Liu, L., Tamer Ozsu, M. (Eds.), *Encyclopedia of Database Systems*. Springer US, pp. 532–538. doi:10.1007/978-0-387-39940-9_565
- Reisberg, D., Hertel, P., 2005. *Memory and emotion*, New York: . ed.
- Renoult, L., Wang, X., Calcagno, V., Prévost, M., Debruille, J.B., 2012. From N400 to N300: variations in the timing of semantic processing with repetition. *Neuroimage* 61, 206–15. doi:10.1016/j.neuroimage.2012.02.069
- Ribeiro, S., Gervasoni, D., Soares, E.S., Zhou, Y., Lin, S.-C., Pantoja, J., Lavine, M., Nicolelis, M. a L., 2004. Long-lasting novelty-induced neuronal reverberation during slow-wave sleep in multiple forebrain areas. *PLoS Biol.* 2, 126–137. doi:10.1371/journal.pbio.0020024
- Roediger, H.L., 2003. Reconsidering implicit memory, in: Bowers, J., Marsolek, C. (Eds.), *Rethinking Implicit Memory*. pp. 3–18.
- Roediger, H.L., 1996. Memory Illusions. *J. Mem. Lang.* 35, 76–100.
- Roland, P.E., Gulyas, B., 1994. Visual imagery and visual representation. *Trends Neurosci.* 17, 281–287.
- Rolls, E.T., 2016. Pattern separation, completion, and categorisation in the hippocampus and neocortex. *Neurobiol. Learn. Mem.* 129, 4–28. doi:10.1016/j.nlm.2015.07.008
- Rolls, E.T., 1996. A theory of hippocampal function in memory. *Hippocampus* 6, 601–620.

- Rolls, E.T., 1989. Functions of neuronal networks in the hippocampus and neocortex in memory, in: Byrne, J., Berry, W. (Eds.), *Neural Models of Plasticity and Theoretical Approaches*. Academic Press: San Diego, pp. 240–265.
- Rosch, E., Mervis, C.B., Gray, W.D., Johnson, D.M., Boyes-Braem, P., 1976. Basic objects in natural categories. *Cogn. Psychol.* 8, 382–439.
- Rosenthal, D.M., 1986. Two concepts of consciousness. *Philos. Stud.* 49, 329–359. doi:10.1007/BF00355521
- Rubin, D., Berntsen, D., 2009. The frequency of voluntary and involuntary autobiographical memories across the life span. *Mem. Cognit.* 37, 679–688. doi:10.3758/37.5.679
- Rubin, D.C., Schulkind, M.D., 1997. The distribution of autobiographical memories across the lifespan. *Mem. Cognit.* 25, 859–866. doi:10.3758/BF03211330
- Rubin, D.C., Wenzel, A.E., 1996. One Hundred Years of Forgetting: A Quantitative Description of Retention. *Psychol. Rev.* 103, 734–760.
- Rugg, M.D., Curran, T., 2007. Event-related potentials and recognition memory. *Trends Cogn. Sci.* 11, 251–7. doi:10.1016/j.tics.2007.04.004
- Ryan, J.D., Althoff, R.R., Whitlow, S., Cohen, N.J., 2000. Amnesia is a deficit in relational memory. *Psychol. Sci.* 11, 454–461.
- Ryan, L., Cox, C., Hayes, S.M., Nadel, L., 2008. Hippocampal activation during episodic and semantic memory retrieval: comparing category production and category cued recall. *Neuropsychologia* 46, 2109–2121.
- Rybash, J., 1999. Aging and autobiographical memory: The long and bumpy road. *J. Adult Dev.* 6, 1–10.
- Sadeh, T., Ozubko, J.D., Winocur, G., Moscovitch, M., 2014. How we forget may depend on how we remember. *Trends Cogn. Sci.* 18, 26–36. doi:10.1016/j.tics.2013.10.008
- Sakai, K., Miyashita, Y., 1991. Neural organization for the long-term memory of paired associates. *Nature* 354, 152–155.
- Sara, S.J., 2000. Retrieval and reconsolidation: toward a neurobiology of remembering. *Learn. Mem.* 7, 73–84.
- Scalaidhe, S., Wilson, F.A., Goldman-Rakic, P.S., 1999. Face-selective neurons during passive viewing and working memory performance of rhesus monkeys: evidence for intrinsic specialization of neuronal coding. *Cereb. cortex* 9, 459–475.
- Schaefer, R.S., Farquhar, J., Blokland, Y., Sadakata, M., Desain, P., 2011. Name that tune: decoding music from the listening brain. *Neuroimage* 56, 843–9. doi:10.1016/j.neuroimage.2010.05.084
- Schatz, C., 1992. The developing brain. *Sci. Am.* 61–67.

- Scherer, K.R., 2009. Emotions are emergent processes: they require a dynamic computational architecture. *Philos. Trans. R. Soc. Lond. B. Biol. Sci.* 364, 3459–74. doi:10.1098/rstb.2009.0141
- Schneider, A.M., Sherman, W., 1968. Amnesia: a function of the temporal relation of footshock to electroconvulsive shock. *Science* (80-.). 159, 219–221.
- Serrano-Pozo, A., Frosch, M.P., Masliah, E., Hyman, B.T., 2011. Neuropathological alterations in Alzheimer disease. *Cold Spring Harb. Perspect. Med.* 1, 1–23. doi:10.1101/cshperspect.a006189
- Serre, T., 2016. Models of visual categorization. *Wiley Interdiscip. Rev. Cogn. Sci.* 7, 197–213. doi:10.1002/wcs.1385
- Shadlen, M.N., Newsome, W.T., 1996. Motion perception: seeing and deciding. *Proc. Natl. Acad. Sci. U. S. A.* 93, 628–633.
- Shadlen, M.N., Shohamy, D., 2016. Decision Making and Sequential Sampling from Memory. *Neuron* 90, 927–939. doi:10.1016/j.neuron.2016.04.036
- Shannon, C.E., 1948. A mathematical theory of communication. *Bell Syst. Tech. J.* 27, 379–423, 623–656.
- Shoham, S., O'Connor, D.H., Segev, R., 2006. How silent is the brain: is there a “dark matter” problem in neuroscience? *J. Comp. Physiol. A* 192, 777–84. doi:10.1007/s00359-006-0117-6
- Silverstein, B.H., Snodgrass, M., Shevrin, H., Kushwaha, R., 2015. P3b, consciousness, and complex unconscious processing. *Cortex* 73, 216–27. doi:10.1016/j.cortex.2015.09.004
- Sirota, A., Csicsvari, J., Buhl, D., 2003. Communication between neocortex and hippocampus 100.
- Smith, S.M., Glenberg, A., Bjork, R.A., 1978. Environmental context and human memory. *Mem. Cognit.* 6, 342–353. doi:10.3758/BF03197465
- Smith, W., 1896. The place of repetition in memory. *Psychol. Rev.* 21–31.
- Snodgrass, J.G., Corwin, J., 1988. Pragmatics of measuring recognition memory: Applications to dementia and amnesia. *J. Exp. Psychol. Gen.* 117, 34–50. doi:10.1037//0096-3445.117.1.34
- Souchay, C., Isingrini, M., Espagnet, L., 2000. Aging, episodic memory feeling-of-knowing, and frontal functioning. *Neuropsychology* 14, 299–309. doi:10.1037/0894-4105.14.2.299
- Souchay, C., Moulin, C.J.A., Clarys, D., Taconnat, L., Isingrini, M., 2007. Diminished episodic memory awareness in older adults: evidence from feeling-of-knowing and recollection. *Conscious. Cogn.* 16, 769–84. doi:10.1016/j.concog.2006.11.002
- Spehn, M.K., Reder, L.M., 2000. The unconscious feeling of knowing: a commentary on Koriat's paper. *Conscious. Cogn.* 9, 187–92; discussion 193–202. doi:10.1006/ccog.2000.0435

- Squire, L.R., 1992. Memory and the hippocampus: A synthesis from findings with rats, monkeys, and humans. *Psychol. Rev.* 99, 195–231. doi:10.1037//0033-295X.99.2.195
- Squire, L.R., 1989. On the course of forgetting in very long-term memory. *J. Exp. Psychol. Learn. Mem. Cogn.* 15, 241–245.
- Squire, L.R., Alvarez, P., 1995. Retrograde amnesia and memory consolidation: a neurobiological perspective. *Curr. Opin. Neurobiol.* 5, 169–177. doi:10.1016/0959-4388(95)80023-9
- Squire, L.R., Cohen, N.J., Nadel, L., 1984. The medial temporal region and memory consolidation: A new hypothesis, in: *Memory Consolidation: Psychobiology of Cognition*. pp. 185–210.
- Squire, L.R., Fox, M.M., 1980. Assessment of remote memory: Validation of the television test by repeated testing during a 7 -year period. *Behav. Res. methods Instrum.* 12, 583–586.
- Squire, L.R., Slater, P.C., 1975. Forgetting in Very Long-Term Memory as Assessed by an Improved Questionnaire Technique. *J. Exp. Psychol. Hum. Learn. Mem.* 104, 50–54.
- Squire, L.R., Stark, C.E.L., Clark, R.E., 2004. The Medial Temporal Lobe. *Annu. Rev. Neurosci.* 27, 279–306. doi:10.1146/annurev.neuro.27.070203.144130
- Squire, L.R., Zola-morgan, S., 1991. The Medial Temporal Lobe Memory System. *Science* (80-.). 253, 1380–1386.
- Standing, L., 1973. Learning 10000 pictures. *Q. J. Exp. Psychol.* 37–41.
- Stokes, M.G., Wolff, M.J., Spaak, E., 2015. Decoding Rich Spatial Information with High Temporal Resolution. *Trends Cogn. Sci.* 19, 636–8. doi:10.1016/j.tics.2015.08.016
- Suzuki, W.A., Amaral, D.G., 1994. Topographic organization of the reciprocal connections between the monkey entorhinal cortex and the perirhinal and parahippocampal cortices. *J. Neurosci.* 14, 1856–1877.
- Swindale, N., 1980. A model for the formation of ocular dominance stripes. *Proc. R. Soc. London. Ser. B, Biol. Sci.* 208, 243–264.
- Taconnat, L., Clarys, D., Vanneste, S., Bouazzaoui, B., Isingrini, M., 2007. Aging and strategic retrieval in a cued-recall test: The role of executive functions and fluid intelligence. *Brain Cogn.* 64, 1–6. doi:10.1016/j.bandc.2006.09.011
- Taconnat, L., Isingrini, M., Taconnat, L., Isingrini, M., 2004. Cognitive Operations in the Generation Effect on a Recall Test: Role of Aging and Divided Attention. *J. Exp. Psychol. Learn. Mem. Cogn.* 30, 827–837. doi:10.1037/0278-7393.30.4.827
- Tang, H., Singer, J., Ison, M.J., Pivazyan, G., Romaine, M., Frias, R., Meller, E., Boulin, A., Carroll, J., Perron, V., Dowcett, S., Arellano, M., Kreiman, G., 2016. Predicting episodic memory formation for movie events. *Sci. Rep.* 6, 30175. doi:10.1038/srep30175

- Tartaglia, E.M., Mongillo, G., Brunel, N., 2015. On the relationship between persistent delay activity, repetition enhancement and priming. *Front. Psychol.* 5, 1–15. doi:10.3389/fpsyg.2014.01590
- Taylor, K.K., Tanaka, K.Z., Reijmers, L.G., Wiltgen, B.J., 2013. Reactivation of neural ensembles during the retrieval of recent and remote memory. *Curr. Biol.* 23, 99–106. doi:10.1016/j.cub.2012.11.019
- Tenenbaum, J.B., Griffiths, T.L., 2002. Generalization, similarity, and Bayesian inference. *Behav. Brain Sci.* 24, 629–640. doi:10.1017/S0140525X01000061
- Thierry, A., Gioanni, Y., Dégénétais, E., Glowinski, J., 2000. Hippocampo-Prefrontal Cortex Pathway : Anatomical and Electrophysiological Characteristics. *Hippocampus* 419, 411–419.
- Thorpe, S., Fize, D., Marlot, C., 1996. Speed of processing in the human visual system. *Nature* 381, 520–522.
- Thorpe, S.J., 2011. Grandmother cells and distributed representations, in: Kriegeskorte, N., Kreiman, G. (Eds.), *Visual Population Codes: Toward a Common Multivariate Framework for Cell Recording and Functional Imaging*. pp. 23–51.
- Thorpe, S.J., 1990. Spike arrival times: A highly efficient coding scheme for neural networks., in: Eckmiller, R., Hartmann, G., Hauske, G. (Eds.), *Parallel Processing in Neural Systems and Computers*. North-Holland Elsevier, pp. 91–94.
- Thorpe, S.J., 1989. Local vs distributed coding. *Intellectica* 2, 3–40.
- Tononi, G., 2008. Consciousness as Integrated Information: a Provisional Manifesto. *Biol. Bull.* 215, 216–242.
- Treves, A., Rolls, E.T., 1992. Computational constraints suggest the need for two distinct input systems to the hippocampal CA3 network. *Hippocampus* 2, 189–200. doi:10.1002/hipo.450020209
- Tsivilis, D., Allan, K., Roberts, J., Williams, N., Downes, J.J., 2015. Old-new ERP effects and remote memories : the late parietal effect is absent as recollection fails whereas the early mid-frontal effect persists as familiarity is retained. *Front. Hum. Neurosci.* 9, 1–11. doi:10.3389/fnhum.2015.00532
- Tsivilis, D., Otten, L., Rugg, M., 2001. Context effects on the neural correlates of recognition memory: an electrophysiological study. *Neuron* 31, 497–505.
- Tulving, E., 1985a. Memory and consciousness. *Can. Psychol. Can.* 26, 1–12.
- Tulving, E., 1985b. How many memory systems are there? *Am. Psychol.* 40, 385–398.
- Tulving, E., 1974. Cue dependent forgetting: when we forget something we once knew it does not necessarily mean that the memory trace has been lost; it may only be inaccessible. *Am. Sci.* 62, 74–82.

- Tulving, E., Donaldson, W. (Eds.), 1972. Episodic and Semantic, in: *Organization of Memory*. pp. 382–402.
- Tulving, E., Madigan, S., 1970. Memory and verbal learning. *Annu. Rev. Psychol.* 21, 437–484.
- Tulving, E., Pearlstone, Z., 1966. Availability versus accessibility of information in memory for words. *J. Verbal Learning Verbal Behav.* 5, 381–391.
- Tulving, E., Schacter, D.L., McLachlan, D.R., Moscovitch, M., 1988. Priming of semantic autobiographical knowledge: A case study of retrograde amnesia. *Brain Cogn.* 8, 3–20. doi:10.1016/0278-2626(88)90035-8
- Tulving, E., Thomson, D., 1973. Encoding specificity and retrieval processes in episodic memory. *Psychol. Rev.* 80, 352–373.
- Turing, A.M., 2007. The Chemical Basis of Morphogenesis. *Philos. Trans. R. Soc. B Biol. Sci.* 237, 37–72.
- Urgolites, Z.J., Wood, J.N., 2013. Visual long-term memory stores high-fidelity representations of observed actions. *Psychol. Sci.* 24, 403–11. doi:10.1177/0956797612457375
- Van Rullen, R., Thorpe, S.J., 2001. Rate coding versus temporal order coding: what the retinal ganglion cells tell the visual cortex. *Neural Comput.* 13, 1255–83.
- Van Strien, J.W., Hagenbeek, R.E., Stam, C.J., Rombouts, S. a R.B., Barkhof, F., 2005. Changes in brain electrical activity during extended continuous word recognition. *Neuroimage* 26, 952–9. doi:10.1016/j.neuroimage.2005.03.003
- VanRullen, R., 2011. Four common conceptual fallacies in mapping the time course of recognition. *Front. Psychol.* 2, 365. doi:10.3389/fpsyg.2011.00365
- Vargha-Khadem, F., Gadian, D.G., Watkins, K.E., Connelly, A., Van Paesschen, W., Mishkin, M., 1997. Differential effects of early hippocampal pathology on episodic and semantic memory. *Science* (80-.). 277, 376–380.
- Verfaellie, M., Cermak, L.S., 1999. Perceptual fluency as a cue for recognition judgments in amnesia. *Neuropsychology* 13, 198–205. doi:10.1037//0894-4105.13.2.198
- Verfaellie, M., Koseff, P., Alexander, M., 2000. Acquisition of novel semantic information in amnesia: effects of lesion location. *Neuropsychologia* 38, 484–492. doi:10.1016/S0028-3932(99)00089-5
- Viard, A., Piolino, P., Lebreton, K., Landeau, B., Young, A., Sayette, V.D. La, Eustache, F., 2007. Hippocampal Activation for Autobiographical Memories over the Entire Lifetime in Healthy Aged Subjects: An fMRI Study. *Cereb. cortex* 17, 2453–2467. doi:10.1093/cercor/bhl153
- Viskontas, I. V., Quiroga, R.Q., Fried, I., 2009. Human medial temporal lobe neurons respond preferentially to personally relevant images. *Proc. Natl. Acad. Sci.* 106, 21329–21334.

- Viswanathan, J., Rémy, F., Bacon-Macé, N., Thorpe, S.J., 2016. Long Term Memory for Noise: Evidence of Robust Encoding of Very Short Temporal Acoustic Patterns. *Front. Neurosci.* 10, 1–11. doi:10.3389/fnins.2016.00490
- Volk, J., 2010. From vibration to manifestation: Assuming our rightful place in creation. *Quester J.* 1–13.
- Von Stein, A., Sarnthein, J., 2000. Different frequencies for different scales of cortical integration: from local gamma to long range alpha/theta synchronization. *Int. J. Psychophysiol.* 38, 301–313. doi:10.1016/S0167-8760(00)00172-0
- Voss, J.L., Baym, C.L., Paller, K.A., 2008. Accurate forced-choice recognition without awareness of memory retrieval. *Learn. Mem.* 15, 454–9. doi:10.1101/lm.971208
- Walker, M.P., Stickgold, R., 2010. Overnight alchemy: sleep-dependent memory evolution. *Nat. Rev. Neurosci.* 11, 218; author reply 218. doi:10.1038/nrn2762-c1
- Wallis, J., Anderson, K., Miller, E., 2001. Single neurons in prefrontal cortex encode abstract rules. *Nature* 411, 953–956.
- Waydo, S., Kraskov, A., Quiroga, R.Q., Fried, I., Koch, C., 2006. Sparse representation in the human medial temporal lobe. *J. Neurosci.* 26, 10232–4. doi:10.1523/JNEUROSCI.2101-06.2006
- Westerberg, C.E., Paller, K.A., Weintraub, S., Mesulam, M.-M., Holdstock, J.S., Mayes, A.R., Reber, P.J., 2006. When memory does not fail: familiarity-based recognition in mild cognitive impairment and Alzheimer's disease. *Neuropsychology* 20, 193–205. doi:10.1037/0894-4105.20.2.193
- Wheeler, M., Stuss, D., Tulving, E., 1997. Toward a theory of episodic memory: the frontal lobes and autonoetic consciousness. *Psychol. Bull.* 121, 331–354.
- Wiggs, C.L., Martin, A., 1998. Properties and mechanisms of perceptual priming. *Curr. Opin. Neurobiol.* 8, 227–233. doi:10.1016/S0959-4388(98)80144-X
- Wiggs, C.L., Weisberg, J., Martin, A., 1998. Neural correlates of semantic and episodic memory retrieval. *Neuropsychologia* 37, 103–118. doi:10.1016/S0028-3932(98)00044-X
- Williams, H.L., Conway, M.A., Moulin, C.J., 2013. Remembering and knowing: using another's subjective report to make inferences about memory strength and subjective experience. *Conscious. Cogn.* 22, 572–88. doi:10.1016/j.concog.2013.03.009
- Wilson, M.A., McNaughton, B.L., 1994. Reactivation of hippocampal ensemble memories during sleep. *Science* (80-.). 265, 676–679.
- Winocur, G., Moscovitch, M., 2011. Memory transformation and systems consolidation. *J. Int. Neuropsychol. Soc.* 17, 766–80. doi:10.1017/S1355617711000683
- Winocur, G., Moscovitch, M., Bontempi, B., 2010. Memory formation and long-term retention in humans and animals: convergence towards a transformation account of hippocampal-neocortical interactions. *Neuropsychologia* 48, 2339–56. doi:10.1016/j.neuropsychologia.2010.04.016

- Wixted, J.T., 2009. Remember/Know judgments in cognitive neuroscience: An illustration of the underrepresented point of view. *Learn. Mem.* 16, 406–412. doi:10.1101/lm.1312809.familiarity
- Wixted, J.T., 2004. The psychology and neuroscience of forgetting. *Annu. Rev. Psychol.* 55, 235–69. doi:10.1146/annurev.psych.55.090902.141555
- Wixted, J.T., Ebbesen, E.B., 1991. On the form of forgetting. *Psychol. Rev.* 2, 409–416.
- Wixted, J.T., Squire, L.R., Jang, Y., Papesh, M.H., Goldinger, S.D., Kuhn, J.R., Smith, K.A., Treiman, D.M., Steinmetz, P.N., 2014. Sparse and distributed coding of episodic memory in neurons of the human hippocampus. *Proc. Natl. Acad. Sci. U. S. A.* 111, 9621–6. doi:10.1073/pnas.1408365111
- Wolk, D.A., Schacter, D.L., Lygizos, M., Sen, N.M., Holcomb, P.J., Daffner, K.R., Budson, A.E., 2006. ERP correlates of recognition memory: effects of retention interval and false alarms. *Brain Res.* 1096, 148–62. doi:10.1016/j.brainres.2006.04.050
- Wu, C., Crouzet, S.M., Thorpe, S.J., Fabre-thorpe, M., 2010. At 120 msec You Can Spot the Animal but You Don ' t Yet Know It ' s a Dog. *J. Cogn. Neurosci.* 27, 141–149. doi:10.1162/jocn
- Yonelinas, A., Aly, M., Wang, W., Koen, J., 2010. Recollection and familiarity: Examining controversial assumptions and new directions. *Hippocampus* 20, 1178–1194. doi:10.1002/hipo.20864.Recollection
- Yonelinas, A.P., 2013. The hippocampus supports high-resolution binding in the service of perception, working memory and long-term memory. *Behav. Brain Res.* October 1, 34–44. doi:10.1016/j.bbr.2013.05.030.The
- Yonelinas, A.P., 2002. The Nature of Recollection and Familiarity: A Review of 30 Years of Research. *J. Mem. Lang.* 46, 441–517. doi:10.1006/jmla.2002.2864
- Zeithamova, D., Schlichting, M.L., Preston, A.R., 2012. The hippocampus and inferential reasoning: building memories to navigate future decisions. *Front. Hum. Neurosci.* 6, 70. doi:10.3389/fnhum.2012.00070

Bringing Very Long-Term Memories back to Life

ABSTRACT

Is it possible to explicitly retrieve sensory information that we have not accessed for decades through recall or familiarity for example? And if so, can we retrieve the information explicitly for example by being able to recall the information or simply having a feeling of familiarity? This question is addressed in this thesis. First, we developed a theoretical model suggesting that it is possible to explicitly retrieve remote and inactive memories. Second, we showed evidence for explicit retrieval of sensory information by testing participants in specific experimental conditions. The maintenance of these memory traces would be visible on Electroencephalography signals. Explicit memory might be more robust to decay than previously thought.

AUTEUR : Christelle LARZABAL

TITRE : Réveiller nos mémoires les plus anciennes

DIRECTEUR DE THÈSE : Simon Thorpe

Thèse soutenue au Centre de Recherche Cerveau et Cognition le : 07/07/2017

RESUMÉ :

Est-il possible de récupérer de manière explicite des informations sensorielles auxquelles nous n'avons pas eu accès depuis des dizaines d'années, sous forme de rappel ou de sentiment de familiarité par exemple? C'est cette question qui a fait l'objet de ce travail de thèse. Dans un premier temps, nous développons un modèle théorique montrant que la récupération explicite de mémoires anciennes laissées dans un état dormant serait possible. Dans un second temps, nous montrons qu'effectivement, en testant des participants dans des conditions expérimentales spécifiques, ces derniers sont capables de récupérer explicitement des informations sensorielles qu'ils pensaient avoir oubliées. Le maintien de ces traces mnésiques serait visible lors d'enregistrements en électroencéphalographie (EEG). Notre mémoire explicite serait donc plus robuste au passage du temps que l'on ne le pensait jusqu'alors.

Mots-clés : Mémoire à long-terme ; Information sensorielle ; Mémoire explicite ; Mémoire inactive ; Rappel ; Familiarité ; Electroencéphalographie

DISCIPLINE : Neuroscience

LABORATOIRE :

*Centre de Recherche Cerveau et Cognition (CerCo) – UMR 5549
Pavillon Baudot-CHU Purpan- BP 25202
31052 Toulouse Cedex*

