



THÈSE DE DOCTORAT DE L'ÉTABLISSEMENT UNIVERSITÉ BOURGOGNE FRANCHE-COMTÉ

PRÉPARÉE À L'UNIVERSITÉ DE BOURGOGNE FRANCHE-COMTÉ

École doctorale n° 554
Environnements-Santé

Doctorat de Neurosciences

par

CORALIE JOUCLA

Électroencéphalographie et machines à vecteurs de support dans le
diagnostic différentiel des pathologies neuropsychiatriques

État des lieux, enjeux et applications

Thèse présentée et soutenue à Besançon, le 7 décembre 2020

Composition du Jury :

SCHRÖDER CARMEN	Professeure, Université de Strasbourg	Présidente du jury
SAMSON SÉVERINE	Professeure, Université de Lille	Rapporteuse
FAKRA ÉRIC	Professeur, Université Jean Monnet, Saint-Étienne	Rapporteur
PERRIN FABIEN	Enseignant-chercheur, Université Claude Bernard, Lyon	Invité
GABRIEL DAMIEN	Chercheur, Université de Franche-Comté, Besançon	Co-encadrant
ORTEGA JUAN-PABLO	Professeur, Université de Saint-Gall, Suisse	Co-encadrant
HAFFEN EMMANUEL	Professeur, Université de Franche-Comté, Besançon	Directeur de thèse



Titre : Électroencéphalographie et machines à vecteurs de support dans le diagnostic différentiel des pathologies neuropsychiatriques

Mots-clés : SVM, EEG, Musique, Alzheimer, Dépression, Coma

Résumé :

Les difficultés diagnostiques sont fréquentes dans les disciplines neurologiques et psychiatriques, où les cliniciens manquent de marqueurs physiologiques pour caractériser certaines maladies, en particulier dans leurs formes débutantes. Depuis quelques années, le recours à l'électroencéphalographie et plus spécifiquement haute résolution (EEG-HR) devient plus courant dans l'exploration clinique des pathologies cérébrales, telles que la dépression, le coma ou les maladies neurodégénératives comme la maladie d'Alzheimer. L'activité cérébrale est aisément enregistrée avec l'EEG, et l'identification de cette activité permet d'isoler les zones du cerveau et le déroulement temporel déclenché par des stimuli spécifiques. L'analyse du signal EEG, et donc son traitement, reste relativement complexe en raison d'un faible rapport signal sur bruit. Dans les études conventionnelles, ces problèmes sont surmontés en faisant la moyenne d'un grand nombre d'essais chez des individus différents. Toutefois, ces méthodes sont difficiles à transposer à une approche clinique où un seul sujet doit être examiné, du fait d'une forte variabilité inter-individuelle. Ce constat appelle le développement d'approches plus robustes pour pallier un rapport signal sur bruit extrêmement défavorable. Les méthodes d'apprentissage automatique, en particulier les Machines à Vecteurs de Support (SVM) ont la

capacité de distinguer des événements EEG qui sont difficilement, voire non distinguables, avec les techniques conventionnelles. Elles offrent dès lors la possibilité de passer de résultats statistiques au niveau du groupe à un résultat personnalisé et par conséquent à un diagnostic donné. Nous proposons dans ce travail de thèse une exploration de la meilleure façon d'extraire les informations pertinentes des enregistrements EEG. Dans un premier temps, nous présentons une revue systématique de la littérature concernant l'application du SVM à l'EEG. Nous mettons en évidence la nécessité d'explorer l'ensemble des paramètres du SVM et d'extraction des caractéristiques du signal pour une performance optimale et reproductible. Nous établissons ensuite si les neuromarqueurs de l'apprentissage de nouvelles chansons sont affectés ou non par l'appréciation musicale via une analyse de groupe, dans le cadre du diagnostic différentiel dépression et maladie d'Alzheimer. Puis nous appliquons le SVM à nos données pour construire des modèles de classification individuels, et nous étudions l'influence de l'optimisation des hyperparamètres. Enfin, nous examinons la fiabilité d'un algorithme génétique pour optimiser l'extraction des caractéristiques avant une classification SVM, pour le diagnostic différentiel des états de conscience pathologiques.

Title: Électroencéphalographie et machines à vecteurs de support dans le diagnostic différentiel des pathologies neuropsychiatriques

Keywords: SVM, EEG, Music, Alzheimer, Depression, Coma

Abstract:

Diagnostic difficulties are common in neurological and psychiatric disciplines, where clinicians lack physiological markers to characterise some diseases, especially in their early forms. In recent years, the use of electroencephalography and more specifically high resolution (EEG-HR) has become more common in the clinical exploration of brain pathologies such as depression, coma or neurodegenerative diseases such as Alzheimer's disease. Brain activity is easily recorded with the EEG, and the identification of this activity makes it possible to single out areas of the brain and the time course triggered by specific stimuli. The analysis of the EEG signal, and therefore its processing, remains relatively complex due to a low signal-to-noise ratio. In conventional studies, these problems are overcome by averaging a large number of trials in different individuals. However, these methods are difficult to transpose to a clinical approach where only one subject has to be examined, due to a high inter-individual variability. This observation calls for the development of more robust approaches to overcome an extremely unfavourable signal-to-noise ratio. Machine learning methods, in

particular Support Vector Machines (SVMs), have the ability to distinguish EEG events that are difficult or even impossible to distinguish with conventional techniques. They thus offer the possibility to go from statistical results at group level to a personalised result and consequently to a given diagnosis. In this thesis work we propose an exploration of how best to extract relevant information from EEG recordings. As a preliminary step, we present a systematic review of the literature regarding the application of SVM to the EEG. We highlight the necessity to explore the full range of SVM parameters and to extract signal characteristics for optimal and reproducible performance. We then establish whether the neuromarkers of new song learning are affected or not by music liking through group analysis, as part of the differential diagnosis of depression and Alzheimer's disease. We then apply SVM to our data to construct individual classification models, and we study the influence of hyperparameters optimisation. Finally, we examine the reliability of a genetic algorithm to optimise features extraction prior to SVM classification, for the differential diagnosis of pathological states of consciousness.

“... tandis que l’individu pris isolément est un puzzle insoluble, il devient, au sein d’une masse, une certitude mathématique. Par exemple, vous ne pouvez jamais prédire ce que fera tel ou tel, mais vous pouvez prévoir comment se comportera un groupe. Les individus varient, mais la moyenne reste constante. Ainsi parle le statisticien.”

Sir Arthur Conan Doyle, Sherlock Holmes : Le Signe Des Quatre

“Pour inspecter notre propre cerveau, nous n’avons que nos propres cerveaux.”

Gordon Rattray Taylor, Le cerveau et ses mystères

“Le silence est nécessaire à la musique mais ne fait pas partie de la musique. Elle s’appuie dessus.”

Robert Bresson, Notes sur le cinématographe

“Vous n’avez pas besoin d’un cerveau pour écouter de la musique.”

Luciano Pavarotti

REMERCIEMENTS

Je voudrais tout d'abord remercier le Pr Emmanuel Haffen, mon directeur, pour m'avoir accueillie dans le laboratoire de Neurosciences avec beaucoup de bienveillance, le Dr Juan-Pablo Ortega et le Dr Damien Gabriel pour avoir accepté de co-encadrer ces travaux de thèse. Je leur suis très reconnaissante pour l'opportunité qu'ils m'ont donnée d'explorer de nouveaux horizons, à la frontière des mathématiques, des neurosciences et de la programmation.

Un merci spécial à Damien, qui de par son accueil pendant mes stages de master a su me donner le goût de la recherche et l'envie d'entreprendre cette grande aventure du doctorat.

Je tiens à remercier particulièrement les Pr Séverine Samson, Eric Fakra et Carmen Schröder pour avoir accepté d'être rapporteurs (pour les deux premiers) et membres du jury. Je suis très honorée par l'intérêt que vous portez à mon travail.

Mes remerciements les plus chaleureux vont à la Dr Lyudmila Grigoryeva, pour m'avoir accueillie au sein de l'université de Konstanz pendant toute une année. Son dynamisme intellectuel, sa curiosité scientifique et son incroyable capacité de travail ont été une véritable source d'inspiration.

Merci à tous mes collègues de recherche, Thibault, d'abord, mon toucan chanteur préféré, et aussi GwenGwen, Aline, Loetitia, Gaëlle et Julie.

J'ai eu l'opportunité grâce à l'Expérimentarium de m'essayer à la vulgarisation scientifique, et je profite de cet espace pour remercier Jérémy, Elodie, Marie, Grégoire et Sindy pour les excellents moments passés avec eux, au musée, au jardin botanique, au collège, à Arles *etc...*

La musique est un élément central dans ce travail de thèse, tout comme dans ma vie. Merci donc à tous les copains du sympho et du Big Band. Merci à Loïc pour tous les projets dingues dans lesquels tu nous embarques, de Cécile Corbel à la source de la Loue, de Moussorgski à Popeye, du classique à la comédie musicale. Merci à Jean-Mi pour mon passage au BB, qui m'a sortie de mes habitudes de jeu classique. Merci à mes copupitres les ptites flûtes, piccolo Bernard, Cécile, Fanny "heySalutCommentCaVaAttendsJ'aiUneSuuuuperBlagueATeRaconter!" et Noémie la petite nouvelle! Merci à mes autres voisins, les hautbois, pour ces moments lunaires où j'ai dû vous dire "c'était là! Joue, joue!".

A Hélène, Cécile et Elodie, merci, votre soutien m'a fait très chaud au cœur.

Merci à l'Allée des Cerisiers, Loïc (encore toi!), Pauline et Yann, votre positive attitude est une bouffée d'oxygène. Merci à toute la bande : Pierre, Aline, Etienne, Saad & Chouquette (j'aurais voulu vous dessiner une poule, puis je me suis rappelé que je ne sais pas dessiner XD), Cédric, Matthieu, Eléonore, Alexis!

Un merci tout particulier à Pierre-Simon, pour ta vision de la musique et de ce qu'elle m'apporte de si précieux : la joie de jouer et de m'exprimer.

Outre les amis musiciens, je voudrais remercier ces musiciens, compositeurs, podcasteurs et autres humoristes, qu'on écoute au casque, dans le creux de l'oreille, et qui, d'une certaine façon, deviennent des proches. C'est donc parti, à fond les ballons, pour un petit tour d'horizon : Luc Lagier, Claude Debussy, Renaud Garcia-Fons, Thomas Enhco, Leon Bridges, Maurice Ravel, Danko Jones, The Cat Empire, Alice Cooper, Snarky Puppy, Pink Floyd, Philip Glass, Ibrahim Maalouf, Karl Jenkins, Kyan Khojandi, Navo, Thomas Wiesel, Guillaume Meurice, Joe Hisaishi, Frédéric Molas et Sébastien Rassiat, Hugues de Courson, Alan Menken, Meshuggah, Tori Kelly et autres John Denver, pour n'en citer que quelques uns.

Musiciens mais aussi mes voisins préférés, Hugo et Mélanie, dont la percu, le violon et le cello m'ont tenu compagnie pendant la rédaction de la thèse, merci pour le violon, les coups de main quand j'étais charrette, nos soirées pizza/jeux et le soutien mutuel pendant le confinement.

Musiciens, encore et toujours, un merci spécial à Eric, Faust mon renard des neiges, Berkan et toute la bande, 12 ans maintenant qu'on se connaît et c'est toujours nimp !

Merci à La Tablée, Louis, Aurore, Cyril, Solveig "Blaguedulion" et Florent Flopinou, votre regard sur le monde, vos réflexions et votre soutien comptent énormément pour moi.

De gros bisous à Hindou et Matthieu, j'ai adoré nos discussions, nos soirées, le tricot ! Je nous souhaite qu'il y en ait beaucoup d'autres.

Merci à ma famille, du fond du cœur, pour la confiance, le soutien et l'amour que vous me donnez. Papa, Maman, je vous aime. Une pensée spéciale pour Raph qui a enfin récupéré le requin au grenier yark yark yark ! et pour Camille qui a la gentillesse de ne pas nous traiter de dingos, alors qu'il y a largement matière. Merci à mes grand-mères, qui ne comprennent pas le concept du doctorat, mais qui soutiennent quand même.

Je voudrais également remercier ma famille d'adoption, Romain et Lydie, Thomas et Lucie, Huguette, Géraldine et Pierre-Olivier. Vous m'avez accueillie dès le début comme un membre à part entière de la famille Meyer, merci pour tous ces bons moments passés tous ensemble !

Quasiment de ma famille aussi (et même de mon ordre et de mon rang, mais on dit des taxons ou des clades ?), à toi, Séb le Bwa, je n'ai pas les mots pour te remercier d'être toujours là, dans les bons moments comme dans les pires, depuis plus de 10 ans maintenant. Yip yip ! RWAAAAAR!!!!

Et toi, Rémi, mon cœur, mon chat, mon encadrant officieux, mon éditeur, mon relecteur, mon soutien chaque jour depuis le début de cette thèse, mon amour, comment te remercier ? Tu sais déjà tout ce que j'ai à te dire.

Enfin, un petit coucou à Marion et Olivier de Mirézain, et à leur chanson "Je remercie... le jury, le public, l'académie etc...", à laquelle je pense très fort en égrenant cette longue liste de remerciements.

SOMMAIRE

I	Introduction	5
1	Contexte médical : diagnostic différentiel en neuropsychiatrie	7
1.1	Diagnostic différentiel maladie d'Alzheimer et dépression	7
1.1.1	Tableaux cliniques	7
1.1.2	Diagnostic différentiel et difficultés	12
1.2	Diagnostic différentiel des troubles de la conscience	14
1.2.1	États de conscience considérés	14
1.2.2	Enjeux spécifiques à ce diagnostic	17
1.2.3	Diagnostic différentiel et difficultés	19
2	L'EEG au service du diagnostic différentiel	21
2.1	Historique : du galvanomètre à l'EEG haute résolution	21
2.2	Générateurs de l'activité électrique captée par EEG	23
2.2.1	Dipôles de courant et organisation spatiale des neurones	24
2.2.2	Potentiels évoqués	26
2.3	Avantages et inconvénients d'un protocole de recherche clinique basé sur l'EEG	27
2.3.1	Pallier des contraintes de neuroimagerie avec l'EEG	27
2.3.2	Limites de l'EEG et <i>machine learning</i> pour le diagnostic clinique	28
3	Une nouvelle approche de l'analyse EEG : le <i>machine learning</i>	31
3.1	Définitions	31
3.1.1	L'ordinateur et l'algorithme	31
3.1.2	L'intelligence artificielle (IA)	31
3.2	Étape 1 : Définition du problème à résoudre	32
3.2.1	Types d'apprentissages	32
3.2.1.1	Apprentissage non supervisé	32
3.2.1.2	Apprentissage supervisé	32
3.2.2	Algorithmes de classification les plus courants	33
3.2.2.1	<i>K-Nearest Neighbours</i>	33

3.2.2.2	<i>Naïve Bayes</i>	33
3.2.2.3	<i>Linear Discriminant Analysis</i>	34
3.2.2.4	<i>Support Vector Machine</i>	34
3.3	Étape 2 : Acquisition des données d'apprentissage et de tests	36
3.3.1	Nécessité d'un grand nombre de données	36
3.3.2	Similarité intraclasse et dissimilarité interclasses	36
3.3.3	Équilibre entre les classes	36
3.4	Étape 3 : Nettoyage des données et extraction des caractéristiques	37
3.4.1	Nettoyage des données	37
3.4.2	Extraction des caractéristiques	38
3.5	Étape 4 : Évaluation de la performance, ajustement du modèle	39
3.5.1	Évaluation de la performance du modèle appris	39
3.5.2	Validation de la performance	40
3.5.3	Ajustement du modèle d'apprentissage	41
3.6	Revue systématique de la littérature	41
3.6.1	Matériel et méthodes	42
3.6.2	Résultats	44
3.6.2.1	Origine des articles	44
3.6.2.2	Résumé de l'étude	44
3.6.2.3	Données utilisées	46
3.6.2.4	Traitement EEG	48
3.6.2.5	Méthodologie SVM	51
3.6.2.6	Résultats obtenus	52
3.6.2.7	Reproductibilité	53
3.6.3	Limitations	53
3.6.4	Recommandations	54
3.6.5	Conclusion	54
II	Dépression et maladie d'Alzheimer : le projet Rihanna	55
4	Modulation de l'encodage et du rappel en mémoire par les stimuli affectifs	63
4.1	Musique, apprentissage et appréciation	63
4.2	Résumé de l'article	64
4.2.1	Introduction	64
4.2.2	Matériel et méthodes	64

4.2.3	Résultats et discussion	64
4.3	Neural signature of musical preference during silence	65
4.3.1	Abstract	65
4.3.2	Introduction	66
4.3.3	Methods	67
4.3.3.1	Participants	67
4.3.3.2	Experimental design	68
4.3.3.3	Data acquisition and analysis	68
4.3.4	Results	70
4.3.4.1	Behavioral analysis	70
4.3.4.2	Time frequency responses	70
4.3.5	Discussion	72
4.3.5.1	Electrophysiological correlates of liking	72
4.3.5.2	Neural sources of EEG activity	73
4.3.5.3	Relationship between liking, exposure, and familiarity	74
4.3.6	Conclusion	74
4.4	Complément à l'article précédent : analyse des ERPs	75
4.4.1	Material and methods : Event Related Potential analysis	75
4.4.2	Results	75
4.4.3	Discussion	75
4.5	Conclusion	75
5	SVM appliqué au projet Rihanna	77
5.1	Résumé de l'article	77
5.1.1	Introduction	77
5.1.2	Matériel et méthodes	77
5.1.3	Résultats et discussion	78
5.2	Machine learning individual differences in the detection of electrophysiological responses of musical familiarity	78
5.2.1	Abstract	78
5.2.2	Introduction	79
5.2.3	Material and methods	80
5.2.3.1	Participants	80
5.2.3.2	Experimental design	80
5.2.3.3	Data acquisition	81

5.2.3.4	Preprocessing	81
5.2.3.5	Features extraction	81
5.2.3.6	SVM classification	82
5.2.3.7	Statistical significance : permutation tests	84
5.2.4	Results	86
5.2.4.1	Classification accuracies	86
5.2.4.2	Kernels	87
5.2.4.3	Features extraction methods	88
5.2.4.4	Regions of interest	88
5.2.4.5	Evolution of hyperparameters	90
5.2.4.6	Density of hyperparameters	92
5.2.4.7	Comparison with standard hyperparameters	94
5.2.4.8	Comparison with best hyperparameters based on test	95
5.2.4.9	Verification of results with randomly generated data	96
5.2.4.10	Results of permutation tests	96
5.2.5	Discussion	98
5.2.5.1	Influence of the selected ROI	98
5.2.5.2	Impact of the features extraction method on performance	99
5.2.5.3	SVM architecture : influence of kernels, hyperparameters and crossvalidation	101
5.2.5.4	Standard hyperparameters	102
5.2.5.5	Randomly generated data	102
5.2.5.6	Statistical analysis of classification accuracies	102
5.3	Conclusion	103
III	États de conscience pathologiques : le projet Cométique	105
6	Le traitement sémantique comme indicateur de conscience résiduelle	107
6.1	Introduction	107
6.2	Matériel et méthodes	109
6.2.1	Protocole expérimental	109
6.2.2	Participants	109
6.2.3	Acquisition des données	109
6.2.4	Pré-traitement	110
6.2.5	Segmentation	111

6.2.6	Analyse temps-fréquence	111
6.2.7	Extraction et optimisation des caractéristiques	111
6.2.7.1	Caractéristiques du signal temporel	111
6.2.7.2	Transformée de Fourier à fenêtre glissante	111
6.2.7.3	Une optimisation individuelle	113
6.2.7.4	Algorithme génétique	113
6.2.8	Support Vector Machine	114
6.3	Résultats	114
6.4	Discussion	122
6.4.1	Potentiels évoqués	122
6.4.1.1	Réponse cérébrale précoce	122
6.4.1.2	Mécanismes attentionnels	122
6.4.1.3	Réponses cognitives	123
6.4.2	Activité fréquentielle : conséquences pour la classification	123
6.4.3	Une forte variabilité inter-individuelle	124
6.4.4	Un surajustement du modèle	124
6.4.4.1	Surajustement dû à l'optimisation des paramètres	124
6.4.4.2	Surajustement dû à des hyperparamètres du SVM trop rigides "standards"	125
6.5	Conclusion	126
IV	Discussion générale et conclusion	127
7	Discussion générale	129
7.1	Contributions de cette thèse	129
7.2	Quelques considérations sur l'IA en médecine	133
7.2.1	Vers une analyse des données plus performante et reproductible	133
7.2.2	Des discours alarmistes dans les médias, mais finalement peu fondés	134
7.2.3	Pourquoi le médecin n'est tout simplement pas remplaçable	134
7.2.4	Mais un optimisme bien présent chez le personnel médical	135
7.2.5	Vers une transformation de l'exercice de la médecine	135
8	Conclusion	137
V	Annexes	179

LISTE DES ARTICLES ET COMMUNICATIONS

Articles

- **Joucla, Coralie**, Gabriel, G., Ortega, J.-H., Vandael, P., Haffen, E., and Grigoryeva, L. **Machine learning individual differences in the detection of electrophysiological responses of musical familiarity.** (in preparation).
- **Joucla, Coralie**, Gabriel, G., Ortega, J.-H., Moulin, T., Haffen, E., and Grigoryeva, L. **On the reliability of machine learning methods for detecting neural responses related to pre-semantic congruency.** (in preparation).
- **Joucla, Coralie**, Gabriel, G., Ortega, J.-H., Haffen, E., and Grigoryeva, L. **Support Vector Machine applied to electroencephalography : a systematic review.** (in preparation).
- **Joucla, Coralie**, Nicolier, M., Giustiniani, J., Brunotte, G., Noiret, N., Monnin, J., Magnin, E., Pazart, L., Moulin, T., Haffen, E., Vandael, P., and Gabriel D. **Evidence for a neural signature of musical preference during silence.** *International Journal of Psychophysiology* (2018), 125 :50-56.

- Gabriel, D., Chabin, T., **Joucla, Coralie**, Bussi re, T., Tarka, A., Galmes, N., Comte, A., Bertrand, G., Giustiniani, J., Haffen, E. **An artistic approach to neurofeedback for emotion regulation.** *NeuroRegulation* (2020), 7 :2.
- Chabin, T., Gabriel, D., Chansophonkul, T., Michelant, L., **Joucla, Coralie**, Haffen, E., Moulin, T., Comte, A., and Pazart, L. **Cortical patterns of pleasurable musical chills revealed by High-Density EEG.** *Frontiers in neuroscience* 14 (2020), 1114.
- Chabin, T., Tio, G., Comte, A., **Joucla, Coralie**, Gabriel, D., and Pazart, L. **The Relevance of a Conductor Competition for the Study of Emotional Synchronization Within and Between Groups in a Natural Musical Setting.** *Frontiers in psychology* 10 (2020), 2954.
- Cabeza, L., Giustiniani, J., Chabin, T., Ramadan, B., **Joucla, Coralie**, Nicolier, M., Pazart, L., Haffen, E., Fellmann, D., Gabriel, D., and Peterschmitt, Y. **Modelling decision-making under uncertainty : A direct comparison study between human and mouse gambling data.** *European neuropsychopharmacology : the journal of the European College of Neuropsychopharmacology* 31 (2020), 58–68.
- Giustiniani, J., **Joucla, Coralie**, Bennabi, D., Nicolier, M., Chabin, T., Masse, C., Trojak, B., Vandael, P., Haffen, E., and Gabriel D. **Behavioral and Electrophysiological Arguments in Favor of a Relationship between Impulsivity, Risk-Taking, and Success on the Iowa Gambling Task.** *Brain Science* (2019), 24 :9(10).

Communications orales et posters

- **Joucla, Coralie, Gabriel, D.**, Ortega, J.-P., Vandel, P., Haffen, E., Grigoryeva, L. (2019). **Un nouveau protocole d'identification de neuromarqueurs de la mémoire musicale basé sur l'ElectroEncéphalographie et la classification par apprentissage machine.** Musique et cerveau, soigner au diapason. (Rouen, France) (communication orale)
- **Joucla, Coralie**, Gabriel D., Grigoryeva L., Ortega J.-P., Haffen E. (2019). **Support Vector Machines applied to EEG : a reliable pipeline for tough problems.** Workshop on Dynamical Systems and Brain-Inspired Information Processing (Konstanz, Germany). (poster)
- **Joucla, Coralie**, Nicolier M., Giustiniani J., Noiret N., Monnin J., Magnin E., Pazart L., Moulin T., Haffen E., Vandel P., Gabriel D. (2019). **Neural signature of musical preference during spontaneous auditory imagery.** NeuroFrance 2019 (Marseille France). (poster)
- **Joucla, Coralie**, Nicolier M., Giustiniani J., Noiret N., Monnin J., Magnin E., Pazart L., Moulin T., Haffen E., Vandel P., Gabriel D. (2018). **Signature neurale de la préférence musicale pendant l'imagerie auditive.** Forum des Jeunes Chercheurs (Besançon, France). (communication orale)
- **Joucla, Coralie**, Gabriel D., Grigoryeva L., Ortega J.-P., Haffen E. (2017). **Personnaliser la détection de signes de conscience au moyen d'un algorithme génétique.** Concours de communication scientifique de la SBB. (Besançon, France). (communication orale)
- **Joucla, Coralie**, Gabriel D., Grigoryeva L., Ortega J.-P., Haffen E. (2017). **A subject personalized frequency domain features selection approach using genetic algorithm to improve SVM classification in EEG N400 paradigm.** Workshop on Dynamical Systems and Brain-Inspired Information Processing (Konstanz, Germany). (poster)
- **Joucla Coralie**, Gabriel D., Grigoryeva L., Ortega J.-P., Haffen E. (2017). **A subject personalized frequency domain features selection approach using genetic algorithm to improve SVM classification in EEG N400 paradigm.** Workshop on Dynamical Systems and Brain-Inspired Information Processing (Konstanz, Germany). (poster)

LISTE DES ABRÉVIATIONS

BCI	<i>Brain-Computer Interface</i>
CIM-11	Classification Internationale des Maladies, 11 ^{ème} version
DSM-IV-TR	Manuel Diagnostique et Statistique des Troubles Mentaux, 4 ^{ème} version, Texte Révisé
EEG-HR	ÉlectroEncéphaloGraphie - Haute Résolution
ERP	<i>Event Related Potentials</i>
ESN	<i>Echo State Network</i>
FFT	<i>Fast Fourier Transform</i>
GA	Algorithme Génétique
GFP	<i>Global Field Power</i>
IA	Intelligence Artificielle
IRMf	Imagerie par Résonance Magnétique fonctionnelle
LIS	<i>Locked-in Syndrome</i>
LPC	<i>Late Positive Component</i>
MCS	<i>Minimally Conscious State</i>
MEG	MagnétoEncéphaloGraphie
PA	Potentiel d'Action
PPS	Potentiel Post-Synaptique
ROI	<i>Region Of Interest</i>
SVM	<i>Support Vector Machine</i>
UWS	<i>Unresponsive Wakefulness Syndrome</i>

INTRODUCTION GÉNÉRALE

Les difficultés diagnostiques sont fréquentes en médecine, en particulier dans les disciplines neurologiques et psychiatriques, où les cliniciens manquent de marqueurs physiologiques pour caractériser certaines maladies, en particulier dans leurs formes débutantes.

Depuis quelques années, le recours à l'électroencéphalographie (EEG) et plus spécifiquement haute résolution (EEG-HR) devient plus courant dans l'exploration clinique des pathologies cérébrales, telles que la dépression, le coma ou les maladies neurodégénératives comme la maladie d'Alzheimer. L'activité cérébrale est facilement enregistrée avec l'EEG, et l'identification de cette activité permet d'isoler les zones du cerveau et le décours temporel déclenché par des stimuli spécifiques. Cependant, l'analyse du signal EEG, et donc son traitement est relativement complexe en raison de multiples parasites : "bruits de fonds", mouvements oculaires, *etc...* et d'une forte variabilité inter-individuelle.

Dans les études conventionnelles, ces problèmes sont le plus souvent surmontés en faisant la moyenne d'un grand nombre d'essais chez des individus différents. L'activité aléatoire disparaît alors que l'activité d'intérêt émerge. Toutefois, ces méthodes sont difficiles à transposer à une approche clinique où un seul sujet doit être examiné : ce constat appelle le développement d'approches plus robustes pour pallier un rapport signal sur bruit extrêmement défavorable.

Récemment, les méthodes d'apprentissage automatique basées sur les mesures EEG, en particulier les Machines à Vecteurs de Support (SVM) ont gagné en popularité dans le domaine médical. Le SVM a la capacité de distinguer des événements qui sont difficilement, voire non distinguables, avec les techniques conventionnelles, avec dès lors la possibilité de passer de résultats statistiques au niveau du groupe à un résultat personnalisé et par voie de conséquence à un diagnostic donné.

Cependant, l'application de l'apprentissage automatique aux données EEG est difficile, en raison du faible rapport signal / bruit et de la grande variabilité inter-individuelle. Nous proposons dans ce travail de thèse une exploration de la meilleure façon d'extraire les informations pertinentes des enregistrements EEG.

Objectifs de la thèse

Ce travail de thèse se développe dans une logique d'amélioration des outils de diagnostic différentiel des pathologies neuro-psychiatriques. Les limites des méthodes actuelles d'investigation clinique, qui ne sont pas satisfaisantes pour les praticiens, mettent en avant le besoin de recourir à des techniques de mesure directe de l'activité cérébrale, telles que l'électroencéphalographie. Les limites de cette méthode de neuroimagerie dans l'exploration des fonctions cognitives de haut niveau, notamment le faible ratio signal sur bruit, restreignent son utilisation à des groupes de sujets, la rendant plus adaptée à

la recherche fondamentale qu'au diagnostic. Il existe donc un besoin de s'orienter vers un niveau d'analyse individuelle de l'EEG, en complément aux stratégies actuelles.

Les objectifs de ce travail s'inscrivent donc dans une démarche exploratoire :

- Le premier objectif est de transposer les méthodes de *machine learning* connues, en particulier le SVM, au traitement du signal EEG. La performance de ces méthodes de traitement du signal sera évaluée sur sujets sains afin de définir si elles permettent un niveau individuel d'analyse des données.
- Le second objectif est d'évaluer dans quelles mesures ces nouvelles méthodes permettent la catégorisation des patients par rapport aux méthodes traditionnelles pour le diagnostic différentiel.
- Un objectif additionnel est la capitalisation des nouveaux savoir-faire acquis, au sein du laboratoire. Bien que des collaborations avec le laboratoire de mathématiques de Besançon aient eu lieu précédemment, ce travail constitue l'incorporation de nouvelles compétences en mathématiques appliquées au sein du laboratoire de neurosciences.

Ce manuscrit comporte quatre parties, la première constitue notre état de l'art, les deux suivantes correspondent aux deux axes cliniques de notre recherche, et enfin, nous discuterons et conclurons sur ce travail dans la dernière partie.

Partie 1 : Introduction

Cette partie inclut trois chapitres. Le premier chapitre établit le contexte médical dans lequel s'inscrivent nos travaux. Nous y décrivons les pathologies dont le diagnostic différentiel nous intéresse, à savoir la maladie d'Alzheimer et la dépression, qui forment notre premier axe de recherche clinique, et les états de conscience pathologiques dans un second axe.

Le chapitre 2 introduit l'électroencéphalographie haute résolution comme méthode de neuroimagerie. Après un rappel historique, nous en présentons les grands principes et les techniques de traitement du signal associées. Un intérêt sera porté aux limitations actuelles de ces méthodes d'analyse et à leur impact sur les investigations cliniques.

Le chapitre 3 présente l'apprentissage machine comme une solution potentielle à ces limites. Nous y exposons les algorithmes les plus couramment utilisés, et nous nous attardons plus spécifiquement sur le Support Vector Machine, qui a retenu notre attention. Nous détaillons le fonctionnement de cette méthode et son implémentation, ainsi que les contraintes associées. Enfin, nous concluons ce chapitre et cette partie par une revue systématique de la littérature concernant l'application du SVM à l'EEG.

Partie 2 : Dépression et maladie d'Alzheimer : le projet Rihanna

Cette partie comporte deux chapitres, tous deux en lien avec notre premier axe clinique : le diagnostic différentiel Alzheimer/dépression. La prévalence de ces deux pathologies, qui touchent fortement la population âgée, est en constante augmentation, mais des symptômes communs, en particulier au niveau des fonctions mnésiques, rendent difficile leur discrimination. Dans ce contexte, un protocole d'imagerie cérébrale en EEG-HR a été

validé sur volontaires sains au Centre d'Investigation Clinique de Besançon afin de mesurer automatiquement des marqueurs neuronaux de l'apprentissage et de la mémoire musicale. Toutefois, la musique est un stimulus pouvant véhiculer de fortes émotions, et il a été établi que l'appréciation ressentie vis-à-vis d'une information est un paramètre facilitateur de l'ancrage en mémoire, ce qui est aimé étant stocké plus rapidement et plus durablement.

Dans le chapitre 4, nous cherchons à savoir si les neuromarqueurs de l'apprentissage sont affectés ou non par l'appréciation musicale chez des sujets sains. Plusieurs mécanismes neuronaux liés à l'appréciation ont été mis en évidence, différents de ceux mis en jeu dans la mémoire. Une réponse neuronale transitoire avec une amplitude modulée par l'appréciation musicale a été constatée.

Dans le chapitre 5, nous nous appliquons à explorer les possibilités offertes par le *machine learning* dans la classification automatique de l'activité cérébrale liée à l'imagerie musicale de chansons connues et inconnues chez des sujets sains. Nous examinons en détail les mécanismes et paramètres du SVM pouvant affecter la performance de classification, outre les données elles-mêmes.

Partie 3 : États de conscience pathologiques : le projet Cométique

Cette partie comporte un unique chapitre, portant sur le diagnostic des différents états de conscience pathologiques, caractérisés par la perte des fonctions de relation. Le questionnement autour de l'état de conscience résiduel des patients est au centre des préoccupations des soignants et des familles, posant la question du prolongement de la vie dans ces conditions et de l'adaptation de la prise en charge médicale de la personne. Étant donné que l'absence de réactions comportementales ne peut être considérée comme la preuve absolue d'une incapacité à comprendre et à vouloir communiquer, nous cherchons donc à explorer directement l'activité neuronale spécifique à certains mécanismes cognitifs, afin de tenter d'établir un pronostic de réveil fiable.

Dans le chapitre 6, l'objectif est d'évaluer la capacité de traitement pré-sémantique du sujet. Le paradigme utilisé repose sur l'utilisation de 120 quatuors de mots. Les trois premiers mots sont congruents et définissent un contexte sémantique, le quatrième étant ou non congruent à ce contexte. En sus d'une analyse traditionnelle des résultats obtenus sur le groupe de sujets sains, nous appliquons le SVM aux données sujet par sujet. Nous optimisons l'extraction des caractéristiques du signal pour la classification à l'aide d'un algorithme génétique.

Partie 4 : Discussion générale et conclusion

Dans cette partie, nous discutons et concluons sur l'ensemble des travaux présentés dans ce manuscrit de thèse. Des perspectives de poursuite de nos recherches sont ensuite proposées. Pour terminer, nous réfléchissons sur la place du *machine learning* et de l'intelligence artificielle d'une manière plus générale dans la médecine.



INTRODUCTION

CONTEXTE MÉDICAL : DIAGNOSTIC DIFFÉRENTIEL EN NEUROPSYCHIATRIE

Les difficultés diagnostiques sont fréquentes en médecine, et ce quel que soit le domaine médical concerné. Cependant, dans certains champs disciplinaires, cette difficulté est exacerbée, en raison de l'absence de critères diagnostiques robustes, de marqueurs biologiques ou physiologiques pathognomoniques, et en raison des intrications fréquentes de plusieurs pathologies, souvent comorbides. Et c'est ce que l'on observe classiquement en psychiatrie et en neurologie : des pathologies frontières, évolutives, pour lesquelles parfois seule l'épreuve du temps aidera à faire la part des choses.

Or le temps est un facteur crucial pour la limitation des risques évolutifs et la mise en place de stratégies thérapeutiques efficaces dès les premiers stades des maladies. Ainsi, de plus en plus, les cliniciens, tant neurologues que psychiatres, questionnent les neurosciences cliniques afin de trouver des outils pertinents pour répondre à ces impératifs. La neuroimagerie est l'un de ces outils qui grâce à l'exploration non invasive et en temps réel du fonctionnement du cerveau pourrait à terme apporter des réponses.

1.1/ DIAGNOSTIC DIFFÉRENTIEL MALADIE D'ALZHEIMER ET DÉPRESSION

1.1.1/ TABLEAUX CLINIQUES

Alzheimer

La maladie d'Alzheimer est la forme de démence la plus fréquente chez la population âgée. Plus de 900 000 personnes sont atteintes par cette maladie en France, qui concerne, en plus des malades, plus de trois millions de personnes (patients et aidants)¹ [1].

Le terme "démence" fait référence à un ensemble de troubles cognitifs résultant de l'endommagement ou de la mort de cellules cérébrales (Tableau 1.1). Par définition, elle provoque un déclin d'au minimum deux de ces quatre fonctions cognitives essentielles : mémoire, capacité langagière, capacité de planification et de jugement, et capacité de traitement des informations visuelles. Le terme démence n'est utilisé que lorsque le déclin

1. Fondation Alzheimer : <https://alzheimer-recherche.org/la-maladie-alzheimer/quest-maladie-dalzheimer/definition-et-chiffres/>

est suffisamment sévère pour avoir un impact sur la vie quotidienne.

Identifiée en 1907 par le psychiatre et neuropathologiste Aloïs Alzheimer (1864-1915), les plaques amyloïdes et les dégénérescences neurofibrillaires sont les marqueurs caractéristiques de la pathologie éponyme. Ces lésions furent pour la première fois décrites lors de la dissection du cerveau d'une patiente présentant des troubles de la mémoire, une désorientation et des hallucinations. Toutefois, celles-ci n'étant identifiables qu'à l'autopsie, le diagnostic de la maladie d'Alzheimer reste pour l'heure un diagnostic d'exclusion.

Critères diagnostiques DSM-IV-TR de la démence de type Alzheimer
<p>A. Apparition de déficits cognitifs multiples comme en témoignent à la fois :</p> <ol style="list-style-type: none"> 1. Altération de la mémoire (altération de la capacité à apprendre des informations nouvelles ou se rappeler les informations apprises ultérieurement) 2. Une ou plusieurs des perturbations cognitives suivantes : <ul style="list-style-type: none"> • Aphasie (perturbation du langage) • Apraxie (altération de la capacité à réaliser une activité motrice malgré des fonctions motrices intactes) • Agnosie (impossibilité à reconnaître ou identifier des objets malgré des fonctions sensorielles préservées) • Perturbation des fonctions exécutives (faire des projets, organiser, ordonner dans le temps, avoir une pensée abstraite)
<p>B. Les déficits cognitifs sont à l'origine d'une altération significative du fonctionnement social ou professionnel et représentent un déclin significatif par rapport au niveau de fonctionnement antérieur</p>
<p>C. L'évolution est caractérisée par un début progressif et un déclin cognitif continu</p>
<p>D. Les déficits cognitifs des critères A1 et A2 ne sont pas dus à :</p> <ol style="list-style-type: none"> 1. D'autres affections du SNC qui peuvent entraîner des déficits progressifs de la mémoire et du fonctionnement cognitif (ex : maladie cérébro-vasculaire, hématome sous-dural...) 2. Des affections générales pouvant entraîner une démence (ex : hypothyroïdie, carence en vitamine B12...) 3. Des affections induites par une substance
<p>E. Les déficits ne surviennent pas de façon exclusive au cours de l'évolution d'un delirium</p>
<p>F. La perturbation n'est pas mieux expliquée par un trouble de l'axe I (ex : schizophrénie, épisode dépressif majeur...)</p>

Tableau 1.1 – Critères diagnostiques de la démence de type Alzheimer. Le début de la démence Alzheimer implique un déclin cognitif progressif et continu. Son diagnostic est un diagnostic d'exclusion. Tableau tiré du Manuel diagnostique et statistique des troubles mentaux IV (texte révisé) [2].

En stade précoce, l'hippocampe (Figure 1.1) et le cortex entorhinal sont les premières régions lésées. Intégrées au système limbique, dans la face interne du lobe temporal médian, ces deux structures fonctionnent en tandem dans les processus de souvenir et d'apprentissage : les stimuli à destination de l'hippocampe passent par l'encodage du cortex entorhinal via la voie perforante. Celui-ci agit donc comme une véritable interface entre le néocortex ainsi que différents noyaux sous-corticaux et l'hippocampe, permettant l'intégration des stimuli dans la mémoire à court terme et la consolidation de la mémoire

à long terme [3, 4].

Des lésions de ce couple entraînent un déficit d'encodage des informations nouvelles en mémoire déclarative et épisodique, se manifestant par un déficit dans l'apprentissage, les fonctions exécutives et dans l'orientation spatio-temporelle [5]. Cependant, attribuer à l'hippocampe un rôle privilégié et unique dans la mémoire épisodique et autobiographique ne suppose pas que le souvenir y soit stocké dans son intégralité. L'activation des circuits para-hippocampiques joue un rôle déterminant dans le codage et la consolidation des épisodes vécus.

Cette consolidation concerne l'ensemble du circuit [mémoire sémantique, mémoire de travail, métamémoire, mémoire à long terme], si bien que des régions extra-hippocampiques pourront porter des marques plus fortes à long terme, comme si la trace des événements avait migré [6, 7].

Ainsi, l'oubli à mesure des patients à lésion hippocampique peut s'accompagner d'une forme limitée d'amnésie rétrograde qui empêche l'accès aux souvenirs constitués quelques mois avant le début du processus pathologique affectant l'hippocampe.

Ce phénomène, selon lequel les souvenirs plus anciens restent accessibles, exprime alors le fait que les connaissances acquises pendant cette période récente n'ont pas été suffisamment consolidés pour pouvoir être réactivées sur sollicitation directe, sans une contribution de l'hippocampe qui les assemblerait dans un registre actif [8, 9].

Les principales zones cérébrales associées à l'hippocampe et touchées par la neurodégénérescence sont les suivantes :

- Le néocortex, la couche la plus externe du cortex, représentant 90% de celui-ci et 76% du cerveau, est particulièrement touché. Il est divisé selon les sutures crânielles en lobes frontaux, pariétaux, temporaux et occipitaux, recouvrant chacun un ensemble de fonctions cérébrales dédiées aux processus cognitifs les plus sophistiqués. Tandis que les aires plus internes du cortex resteront relativement intactes, préservant l'activité sensorielle, la dégénérescence du néocortex entraînera à mesure un déclin, des changements de fonctionnement de personnalité. La progression de la destruction neuronale dans le cortex associatif temporo-pariéto-frontal, entraînera à mesure des troubles du langage, une désorientation dans le temps et l'espace, un affaiblissement du jugement, des changements d'humeur et de comportement – tels qu'une agressivité, un repli sur soi –, une dégradation de la mémoire sémantique, et enfin une démence [2, 10, 11].
- L'amygdale, située dans la région antéro-interne du lobe temporal, en avant de l'hippocampe (Figure 1.1), est un noyau pair impliqué dans l'évaluation de la valence émotionnelle des stimuli, notamment dans les réponses à la peur ou à l'anxiété. Elle fonctionne en coordination avec l'hippocampe et le cortex entorhinal lors de la mise en mémoire d'événements à forte connotation émotionnelle : l'amygdale va non seulement renforcer l'encodage [12], mais aussi moduler la consolidation des souvenirs épisodiques [13]. En effet, après avoir été encodée, la trace mnésique doit être consolidée pour ne pas disparaître, et l'amygdale va permettre une rétention plus forte et plus durable des souvenirs émotionnellement chargés, par rapport aux souvenirs d'événements neutres. La détérioration de cette structure cérébrale, et la dégradation de sa qualité d'interaction avec l'hippocampe vont donc conduire à une perte de la capacité à apprécier la signification émotionnelle d'une nouvelle expérience, et participer au phénomène d'amnésie antérograde observé chez les

patients atteints de la maladie d'Alzheimer.

— Trois noyaux présents dans la formation réticulée sont également impliqués :

- Le noyau basal de Meynert, associé aux stimuli activateurs, tant positifs que négatifs [14], est également impliqué dans le maintien de l'attention [15], l'apprentissage et le rappel dans la mémoire à long terme [16]. Sa dégénérescence semble étroitement liée à l'apparition des symptômes de démence [17].
- Le *Locus Coeruleus*, lié à de nombreuses fonctions différentes de par ses projections exhaustives, a une influence sur les rythmes circadiens, l'attention et la mémoire, le comportement et les émotions [18]. La destruction de ce noyau – jusqu'à 80% [19] – a été associée à une accélération de la progression de la pathologie dans des modèles murins [20, 21].
- Le noyau du Raphé participe à la régulation de l'humeur de part sa très forte proportion en neurones sérotoninergiques [22]. La dégénérescence de cette entité cérébrale participe aux troubles de l'humeur constatés dans la maladie d'Alzheimer.

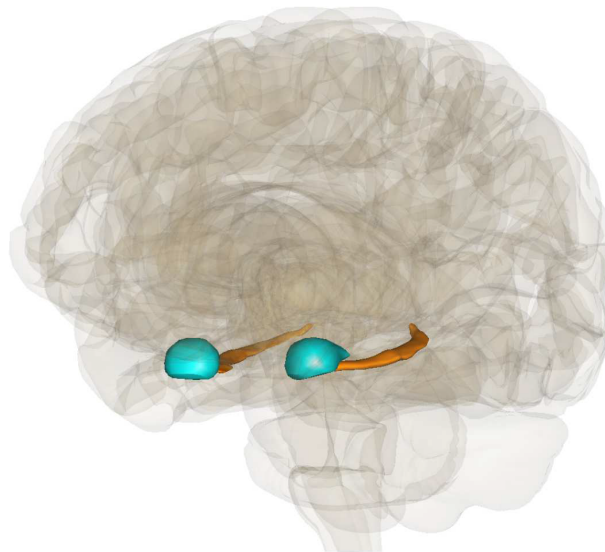


FIGURE 1.1 – En bleu, l'amygdale fonctionne en lien avec l'hippocampe et le cortex entorhinal lors de la mise en mémoire d'événements à connotation émotionnelle. Elle est impliquée tant dans la maladie d'Alzheimer que dans la dépression. En orange, l'hippocampe est la première région lésée lors de la neurodégénérescence induite par la maladie d'Alzheimer. Image générée à partir d'Anatomography, site web administré par Life Science Databases. <https://lifesciencedb.jp/bp3d/#>

Dépression

La dépression est un trouble de l'humeur qui se distingue par une tristesse pathologique, une perte d'intérêt et du plaisir (ou anhédonie), un ralentissement psychomoteur et des troubles somatiques, en particulier des fonctions instinctuelles (sommeil, alimentation, etc...), et des troubles cognitifs (Tableau 1.2) [2].

Il s'agit de l'affection la plus communément observée chez le sujet âgé. Il est en effet estimé que 10 à 45% des adultes de plus de 65 ans souffrent de ce trouble majeur de l'humeur [23]. Bien qu'ils soient très fréquents chez le sujet âgé, les troubles dépressifs restent souvent méconnus (60% à 70% des sujets dépressifs), et ne sont que peu traités. En avril 2017, la dépression a été annoncée comme étant devenue la première cause d'invalidité liée à la santé par l'Organisation Mondiale de la Santé. On estime que plus de 322 millions de personnes souffrent de dépression, soit 4,4% de la population mondiale [24].

La dépression entraîne une altération importante du fonctionnement quotidien. Le fonctionnement social diminue en corrélation avec l'augmentation de la gravité de la dépression. En général, on a constaté que la dépression augmentait le risque de handicap social de 23 fois par rapport à la population générale.

Elle est caractérisée par des difficultés au niveau du mécanisme de rappel des informations mais pas de l'encodage. Ce mécanisme est altéré tant pour le passé ancien que récent, en particulier lorsque la restitution nécessite une recherche active. En outre, a été observée la survenue de souvenirs intrusifs de valence négative, tandis que le rappel d'événements positifs spécifiques est altéré au profit de souvenirs non spécifiques répétés présentant un contenu émotionnel moins riche (phénomène de surgénéralisation) [25] : en réponse à un mot-indice, les patients déprimés produisent moins de souvenirs spécifiques, c'est-à-dire se rapportant à un événement survenu une seule fois et ayant duré moins d'une journée. Le rappel autobiographique des patients déprimés s'explique surtout par un excès de souvenirs catégoriels, c'est à dire se référant à des événements survenus plus d'une fois. Cette surgénéralisation est associée à un moins bon pronostic à court terme de l'épisode dépressif, notamment chez les patients les plus âgés.

Les souvenirs intrusifs sont caractérisés par leur rappel involontaire, rapide, sans effort cognitif et sans lien avec le contexte. Jusqu'à 85% des patients déprimés rapportent des souvenirs intrusifs. En général liés à des événements de vie négatifs, il s'agit de souvenirs incontrôlables, détaillés et associés à un sentiment de reviviscence. Par rapport aux souvenirs intrusifs des sujets sains, ils sont généralement plus vivaces, associés à une détresse émotionnelle plus intense et engendrent davantage d'interférences.

L'état dépressif est susceptible d'entraîner l'apparition de modifications neurocognitives, telles que des baisses de performances touchant l'attention, les fonctions exécutives et la mémoire de rappel [26–29]. Le déficit cognitif peut atteindre une intensité telle qu'il peut remplir les critères diagnostiques d'un état démentiel [30] même si, le plus souvent, le déficit observé dans la maladie d'Alzheimer semble plus important [24].

L'hippocampe joue un rôle essentiel dans la formation de la mémoire, en tant que partie du circuit impliqué dans le traitement de l'information et la création de souvenirs émotionnels et déclaratifs. Le volume de l'hippocampe est diminué chez les patients souffrant de dépression, en particulier en cas d'épisodes récurrents ou chroniques ou d'antécédents de traumatismes. La neurogenèse altérée a été invoquée pour expliquer cette constatation, car l'augmentation de la sécrétion de glucocorticoïdes due à un stress prolongé est particulièrement neurotoxique pour les neurones de l'hippocampe.

Dans la dépression, le dysfonctionnement du réseau impliquant ces circuits limbiques - corticaux - avec une sous-activité dans les régions corticales (cortex préfrontal dorso-latéral, cortex cingulaire antérieur dorsal) et une suractivité dans le limbique (amygdale, hippocampe, *nucleus accumbens*), paralimbique (cortex cingulaire antérieur ventral, cortex préfrontal ventromédial) et les régions corticales intégratives - sont supposées en-

Critères diagnostiques de l'épisode dépressif majeur selon le DSM-IV-TR
<p>A. Au moins cinq des symptômes suivants doivent être présents durant une période de 2 semaines minimum et avoir représenté un changement par rapport au fonctionnement antérieur. Au moins un des symptômes est soit une humeur dépressive soit une perte d'intérêt ou de plaisir.</p> <ol style="list-style-type: none"> 1. Humeur dépressive présente pratiquement toute la journée, presque tous les jours, signalée par le sujet (ex : pleure). N.B. : Eventuellement irritabilité 2. Diminution marquée de l'intérêt ou du plaisir pour toutes ou presque toutes les activités, pratiquement toute la journée, presque tous les jours (signalée par le sujet ou observée par les autres). 3. Perte ou gain de poids significatif en l'absence de régime (ex : modification du poids corporel en un mois excédant 5%), diminution ou augmentation de l'appétit presque tous les jours. N.B. : Chez l'enfant, prendre en compte l'absence de l'augmentation de poids attendue. 4. Insomnie ou hypersomnie presque tous les jours. 5. Agitation ou ralentissement psychomoteur presque tous les jours (constaté par les autres, non limité à un sentiment subjectif de fébrilité ou de ralentissement intérieur). 6. Fatigue ou perte d'énergie presque tous les jours. 7. Sentiment de dévalorisation ou de culpabilité excessive ou inappropriée (qui peut être délirante) presque tous les jours (pas seulement se faire grief ou se sentir coupable d'être malade). 8. Diminution de l'aptitude à penser ou à se concentrer ou indécision presque tous les jours (signalée par le sujet ou observée par les autres). 9. Pensées de mort récurrentes (pas seulement une peur de mourir), idées suicidaires récurrentes sans plan précis ou tentative de suicide ou plan précis pour se suicider.
B. Les symptômes ne répondent pas aux critères d'épisode mixte.
C. Les symptômes induisent une souffrance cliniquement significative ou une altération de fonctionnement social, professionnel ou dans d'autres domaines importants.
D. Les symptômes ne sont pas imputables aux effets physiologiques directs d'une substance (ex : une substance donnant lieu à abus, un médicament) ou d'une affection générale (ex : hypothyroïdie).
E. Les symptômes ne sont pas mieux expliqués par un Deuil.

Tableau 1.2 – Critères diagnostiques de l'épisode dépressif majeur. La caractéristique principale de l'épisode dépressif est une humeur dépressive ou une perte d'intérêt ou de plaisir pour presque toutes les activités pendant au moins deux semaines. Tableau tiré du Manuel diagnostique et statistique des troubles mentaux IV (texte révisé) [2].

traîner des symptômes d'humeur et de comportement et une dysrégulation hypothalamique associée.

1.1.2/ DIAGNOSTIC DIFFÉRENTIEL ET DIFFICULTÉS

A l'heure actuelle, l'évaluation cognitive des patients consultant en psychiatrie repose principalement sur l'utilisation de bilans neuropsychologiques.

Deux systèmes internationaux de classification des pathologies neuropsychiatriques, la Classification Internationale des Maladies, de l'Organisation Mondiale de la Santé, onzième version (CIM-11) [31] et le Manuel de Diagnostic Statistique, quatrième version (DSM-IV-TR) [2], sont principalement réalisés pour poser un diagnostic de dépression. Différents questionnaires issus de ces classifications permettent le dépistage, l'évaluation de la sévérité du trouble dépressif, et le suivi de la pathologie.

Dans le cas de la maladie d'Alzheimer, les symptômes de déficit mnésique caractéristiques apparaissent dès le stade léger de la pathologie et sont le plus souvent identifiés par la famille ou l'entourage du patient. Le *Mini Mental State Examination* [32] est l'un des tests le plus souvent utilisés pour le dépistage de la maladie d'Alzheimer. Il permet d'évaluer l'orientation spatio-temporelle, la mémoire, l'attention, le langage et les praxies constructives. Le score, évalué de 0 à 30, permet d'évaluer la progression de la pathologie, un nombre élevé indiquant un stade sévère.

Néanmoins, ces bilans restent insuffisants pour poser un diagnostic dans les stades les plus précoces [33, 34]. En effet, il existe un certain recouvrement sémiologique de ces deux pathologies qui rend le diagnostic différentiel particulièrement délicat.

Maladie d'Alzheimer	Dépression
Lésions neurodégénérantes	Tristesse intense, longue durée
Déficit encodage épisodique	Déficit rappel actif

Symptômes communs au stade précoce	
Troubles psychocomportementaux :	Pseudo-démence :
Apathie Anxiété	Désorientation, troubles attentionnels Maladresse
=> Troubles dépressifs ?	=> Apraxie ?

Tableau 1.3 – Difficultés dans le diagnostic différentiel maladie d'Alzheimer et dépression chez le sujet âgé. Des symptômes de la maladie d'Alzheimer (en orange) peuvent apparaître dans le tableau clinique du patient dépressif (en bleu) et réciproquement.

Chez la personne âgée, la pathologie dépressive se présente fréquemment de manière atypique, avec des troubles cognitifs pouvant aller jusqu'à la faire apparaître comme une "pseudo-démence". Lors de ce phénomène, le patient présente des troubles mnésiques, une désorientation, des troubles attentionnels et une maladresse rappelant à tort l'apraxie [35]. Il semblerait que la dépression chez le sujet âgé puisse présenter des symptômes différents de ceux reportés chez le jeune adulte, avec des sous-types de troubles liés au vieillissement. Ainsi les troubles cognitifs constatés ne correspondent pas toujours à la forme clinique de dépression décrite dans le DSM-IV-TR et dans la CIM-11 (Tableau 1.3).

La maladie d'Alzheimer, quant à elle, apparaît souvent avec des troubles psychocomportementaux associés au syndrome amnésique hippocampique, tels que l'apathie et une symptomatologie anxieuse et dépressive à tous les stades, faisant évoquer à tort des troubles dépressifs [36].

En complément de l'examen neuropsychologique, un examen d'Imagerie par Résonance Magnétique (IRM) est aussi souvent effectué dans un premier temps afin de rechercher une altération morphologique correspondant à une pathologie identifiée et/ou de préciser l'origine et la nature des troubles cognitifs. Cependant, ces deux examens ne permettent pas de poser le diagnostic de maladie d'Alzheimer et/ou de dépression avec certitude. Le traitement antidépresseur d'épreuve, où le patient reçoit un traitement antidépresseur, reste dans certains cas le seul moyen de discriminer ces deux pathologies comorbides [37] ; toutefois ce test présente également ses limites, car tous les patients ne répondent pas aux antidépresseurs [38].

Un rapport ministériel publié en France en 2000 rapporte que près de 50% des patients testés ne pourraient être discriminés de manière certaine, et que seuls 20% des sujets diagnostiqués recevraient une prise en charge spécifique [39]. Dans un tel contexte, la définition de nouveaux outils de diagnostic et la recherche de neuromarqueurs permettant d'affiner les bilans effectués actuellement apparaît comme indispensable [40]. De nouvelles méthodes de discrimination objectives de ces deux populations permettraient une prise en charge adaptée du patient au plus tôt, améliorant ainsi le pronostic, puisque l'approche thérapeutique n'est pas la même [41, 42].

1.2/ DIAGNOSTIC DIFFÉRENTIEL DES TROUBLES DE LA CONSCIENCE

Le domaine d'étude des états de conscience altérée est apparu durant les années 1950. En effet, l'état de la technologie médicale auparavant ne permettait pas aux patients souffrant de graves atteintes cérébrales de survivre, en particulier lorsque les zones responsables des fonctions vitales étaient touchées. L'émergence de cette discipline relativement récente est liée à l'apparition de la respiration artificielle et au développement des soins intensifs, qui ont permis la dissociation des fonctions cardiaques, respiratoires et cérébrales. De nombreux patients vont ainsi voir leur vie prolongée, mais dans des états cliniques critiques auxquels jusque-là les membres du personnel médical n'avaient été que très peu confrontés [43]. En 2002, le nombre de patients en état prolongé de conscience altérée en France était estimé à environ mille cinq cent, soit une prévalence d'environ vingt-cinq cas pour un million d'habitants², tandis qu'il serait de quarante-six pour un million aux Etats-Unis, et de quatorze pour un million au Royaume-Uni [44].

1.2.1/ ÉTATS DE CONSCIENCE CONSIDÉRÉS

Coma

C'est en 1966 que Plum et Posner posent pour la première fois les bases du diagnostic d'une nouvelle entité clinique : le coma [45]. Le coma, dont l'étymologie signifie "profond sommeil", traduit une altération pathologique de la conscience d'une personne. Le terme "conscience" peut désigner dans ce cas deux concepts distincts qui sont la conscience de soi ("qui je suis", "comment je me sens" *etc...*) et la conscience de son environnement c'est-à-dire, la perception qu'a la personne du monde l'entourant. La conscience de soi n'est pas évaluable par un observateur externe, puisque seul le sujet lui-même sait ce qu'il ressent. L'éveil, repérable à l'ouverture spontanée des yeux, est également pris en compte pour déterminer le niveau de conscience d'une personne.

Cet état n'est que temporaire, et peut évoluer vers différentes issues : certains patients récupèrent très rapidement, avec ou sans séquelles, dont la gravité est variable, alors que d'autres voient leur santé se dégrader, et leur état évoluer vers la mort cérébrale. Enfin, la plus grande majorité des patients en coma récupèrent progressivement de la conscience, en passant par d'autres états de conscience altérée que nous allons définir à présent (Figure 1.2).

2. Circulaire DHOS/02/DGS/SD5D/DGAS n° 2002-288 du 3 mai 2002 relative à la création d'unités de soins dédiées aux personnes en état végétatif chronique ou en état pauci-relationnel, <https://solidarites->

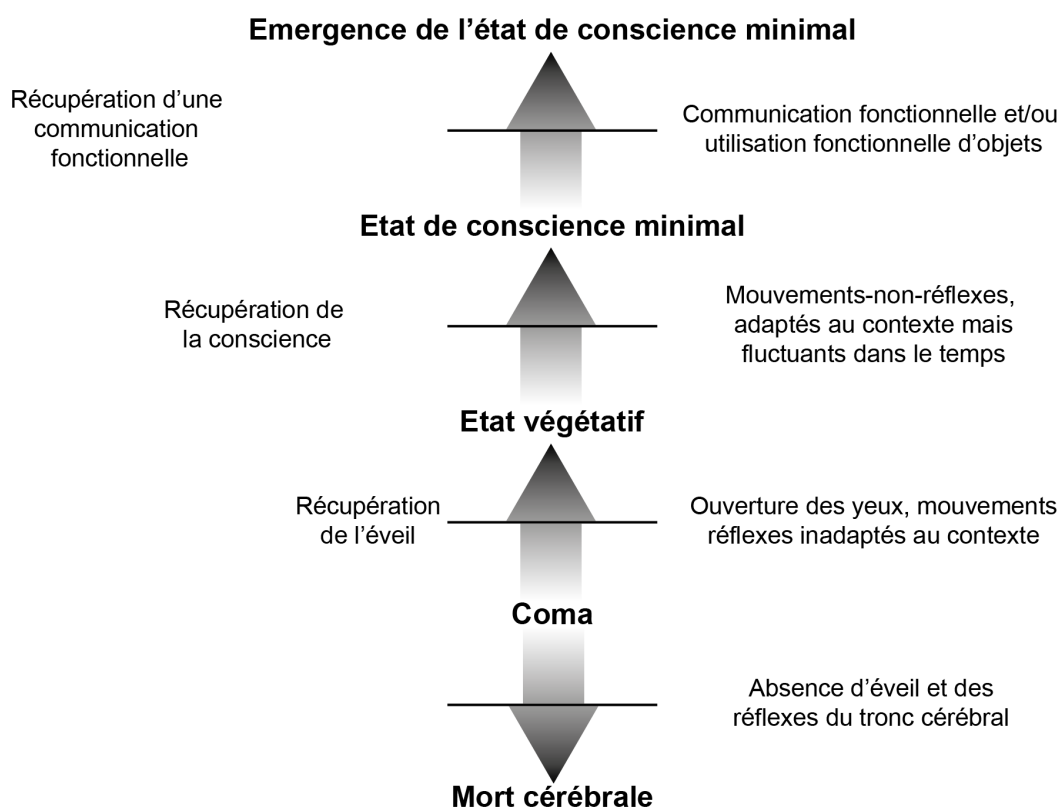


FIGURE 1.2 – Les différentes étapes de récupération. Un patient en état végétatif (UWS) va généralement évoluer vers un état de conscience minimale (MCS), puis éventuellement émerger de cet état vers une récupération complète ou avec séquelles neurologiques ou psychiques. Tiré de Bruno *et al.* (2011). [46]

Mort cérébrale

Les progrès cliniques et technologiques et le maintien en vie de personnes en états de conscience limités ont conduit à une caractérisation alternative de la mort, passant d'une définition cardio-respiratoire, où l'arrêt cardiaque signifiait la mort du patient, à une définition neuro-centrée, où la mort est à présent définie comme la perte irréversible de tous les réflexes du tronc cérébral [47].

Dans cet état irréversible, le patient, apnéique et parfaitement aréactif, n'est gardé en vie que par la grâce des machines qui maintiennent ses fonctions cardiaques et respiratoires. Ce diagnostic fait l'objet d'une procédure standardisée et complexe afin de s'assurer de l'irréversibilité des lésions du patient. L'absence de réflexes du tronc cérébral tels que les réflexes cornéens, pupillaires, nauséux, de toux et oculo-vestibulaires doit être démontrée par une batterie de tests, tandis que l'absence totale d'activité cérébrale doit être confirmée par la réalisation de deux ÉlectroEncéphaloGrammes (EEG) à au moins quatre heures d'intervalle. Ces tracés EEG doivent être déclarés plats par deux médecins différents pour que le diagnostic de mort cérébrale soit validé. Un test d'apnée est également réalisé : la ventilation artificielle est stoppée, et si la respiration ne reprend pas spontanément, la mort cérébrale est ainsi confirmée.

Etat végétatif chronique/syndrome d'éveil non répondant

Dénommé "état végétatif chronique" jusqu'à il y a quelques années, le syndrome d'éveil non répondant (*Unresponsive Wakefulness Syndrome* ou UWS) est l'état vers lequel évoluent rapidement la majorité des comas. Dans ce cas, ces patients ne sont pas maintenus en vie artificiellement et présentent des épisodes d'ouverture spontanée des yeux, le cycle veille-sommeil est préservé ainsi que les fonctions autonomes telles que les fonctions cardiaques et respiratoires.

Ces patients peuvent émettre des sons, voire prononcer des mots, sans que ceux-ci ne soient corrélés au contexte situationnel. Ils peuvent en outre avoir des comportements particuliers tels que des rires ou des pleurs, mais sans raison apparente. Ces comportements ne sont pas volontaires et ne témoignent pas de décisions conscientes mais plutôt d'activité réflexe. Ils ne peuvent donc pas communiquer.

Cet état peut être permanent ou évoluer à son tour vers l'état de conscience minimal. Il est généralement admis que plus la durée de l'UWS est faible, plus les chances de récupération sont élevées.

État de conscience minimal

Comme l'indique son nom, les patients en état de conscience minimal (*Minimally Conscious State* ou MCS) possèdent un degré "minimal" de conscience. En effet, des réponses aux stimuli environnementaux peuvent être observées de manière reproductible, contrairement à l'UWS. Des comportements de poursuite visuelle et d'obéissance à des ordres simples, voire même de communication élémentaire, par gestes ou verbalement, et des manifestations émotionnelles adaptés au contexte peuvent parfois être constatés. Le diagnostic différentiel entre l'UWS et le MCS est particulièrement crucial en raison des décisions thérapeutiques en découlant.

Locked in syndrome

Le coma peut également parfois évoluer en "*Locked-In Syndrome*" (LIS), ou syndrome d'enferment. Le patient est alors totalement conscient et éveillé, et ses fonctions cognitives sont intactes, mais il souffre d'une tétraplégie et d'une anarthrie qui l'empêchent de bouger et de parler. Les similitudes de cet état sont fortes avec l'UWS, notamment l'incapacité à se mouvoir et l'ouverture spontanée des yeux. Le diagnostic est donc particulièrement délicat à poser, et n'intervient généralement que 2 mois et demi après l'incident déclencheur de l'état. Ce sont les proches qui le plus souvent détectent en premier les signes de conscience (dans 55% des cas), plutôt que l'équipe soignante (23% des cas).

Il existe plusieurs catégories de LIS : généralement, les patients n'ont aucun contrôle musculaire, mais peuvent parfois contrôler leurs mouvements oculaires, voire récupérer certains mouvements volontaires. Dans tous les cas, ces patients sont pleinement conscients de leur corps et de leur environnement, mais ne peuvent que difficilement s'exprimer. De plus, le système nerveux étant intact, ils ressentent le toucher ainsi que la douleur.

Des stratégies de communication peuvent être mises en place : les patients peuvent apprendre à utiliser des ordinateurs contrôlés par les mouvements oculaires, détectés par une caméra. Une récupération, même minime, des mouvements volontaires permet

d'équiper les patients d'appareils sophistiqués afin de leur redonner le plus d'autonomie possible. Ainsi, il est de plus en plus fréquent que des patients atteints de LIS envisagent un retour à domicile.

Bien que le LIS ne soit pas un état de conscience altéré, puisque le sujet reste pleinement conscient, nous avons choisi de le présenter ici, car il constitue l'une des possibles issues du coma.

1.2.2/ ENJEUX SPÉCIFIQUES À CE DIAGNOSTIC

Emotionnels et familiaux

Avoir un proche dans un état de conscience altéré est une épreuve difficile pour la famille et l'entourage. Tout d'abord, le diagnostic différentiel étant complexe, il est compliqué pour les médecins de pronostiquer l'issue, favorable ou non, pour le patient, ainsi que les éventuelles séquelles. L'incertitude et l'inquiétude sont donc très présentes chez les proches.

D'autre part, la décision de cesser ou non les soins, décision dont les enjeux éthiques et socio-culturels, que nous détaillerons ci-après, sont très importants, est dans la plupart des cas prise en accord avec la famille, dans une approche centrée sur le patient : il s'agit de prendre une décision en accord avec ce que le patient aurait souhaité [48]. Toutefois, si dans certains cas le patient a fait part à l'écrit de ses *desiderata*, les soignants et les proches se retrouvent le plus souvent à devoir deviner ses préférences. Cet "état de fait" fait peser une très lourde charge, tant éthique qu'émotionnelle, sur les personnes chargées de prendre cette décision.

Socioculturels

Aujourd'hui encore, le terme de "conscience" revêt de multiples facettes. Or, ce sont les divergences dans la façon dont nous envisageons ce concept qui mène le plus souvent aux controverses vis-à-vis de la prise en charge médicale des patients souffrant de troubles de la conscience.

Les états limites de conscience ont mené à l'apparition du concept d'"acharnement thérapeutique", ou "obstination déraisonnable". Ces deux termes désignent l'utilisation de thérapies disproportionnellement lourdes par rapport à l'amélioration de l'état du patient que l'on peut en attendre, tout en sachant que le refus de cet acharnement peut potentiellement conduire à la mort. En France, l'interdiction de l'acharnement thérapeutique a été introduite dans le code de la santé publique par la loi Leonetti de 2005³, et précisée par la loi Claeys-Leonetti de 2016⁴ : "lorsque [les actes médicaux mis en œuvre] apparaissent inutiles, disproportionnés ou lorsqu'ils n'ont d'autre effet que le seul maintien artificiel de la vie, ils peuvent être suspendus ou ne pas être entrepris, conformément à la volonté du patient et, si ce dernier est hors d'état d'exprimer sa volonté, à l'issue d'une procédure collégiale définie par voie réglementaire".

3. Loi n° 2005-370 du 22 avril 2005 relative aux droits des malades et à la fin de vie, JORF n°0028 du 3 février 2016, <https://www.legifrance.gouv.fr/jorf/id/JORFTEXT000031970253/>

4. Loi n° 2016-87 du 2 février 2016 créant de nouveaux droits en faveur des malades et des personnes en fin de vie, JORF n°0028 du 3 février 2016, <https://www.legifrance.gouv.fr/jorf/id/JORFTEXT000031970253/>

Diagnostic différentiel des états de conscience modifiés					
Etat	Etat végétatif	Etat de conscience minimal	Syndrome Locked-in	Coma	Mort confirmée par des tests sur le tronc cérébral
Conscience	Absente	Présente	Présente	Absente	Absente
Cycle veille-sommeil	Présent	Présent	Présent	Absent	Absent
Réponse à la douleur	±	Présent	Présent (dans les yeux uniquement)	±	Absent
Glasgow Coma Score	E4, M1-4, V1-2	E4, M1-5, V1-4	E4, M1, V1	E1, M1-4, V1-2	E1, M1-3, V1
Fonction motrice	Aucun mouvement intentionnel	Comportements verbaux ou moteurs intentionnels cohérents ou incohérents	Mouvements oculaires verticaux volontaires ou clignement des yeux préservés	Pas de mouvement intentionnel	Pas de mouvement, ou uniquement mouvements de réflexes spinaux
Fonction respiratoire	Généralement préservée	Généralement préservée	Généralement préservée	Variable	Absente
Activité EEG	Généralement activité à ondes lentes	Données insuffisantes	Généralement normale	Généralement activité à ondes lentes	Généralement absente
Métabolisme cérébral (PET)	Sévèrement réduit	Données insuffisantes	Légalement réduit	Modérément à sévèrement réduit	Sévèrement réduit ou absent
Pronostic	Variable : si permanent, état végétatif continu ou décès	Variable	Dépend de la cause mais rétablissement complet improbable	Rétablissement, état végétatif ou décès dans les semaines qui suivent	Déjà mort

Tableau 1.4 – Taxonomie britannique des pathologies persistantes majeures de la conscience. Traduit de Zeman (2008). [49]

1.2.3/ DIAGNOSTIC DIFFÉRENTIEL ET DIFFICULTÉS

La difficulté de la question du non-acharnement thérapeutique réside dans la limite avec l'euthanasie, dont l'objectif est de provoquer intentionnellement le décès d'une personne, en principe atteint d'une maladie incurable qui lui inflige des souffrances intolérables, particulièrement par un médecin ou sous son contrôle. Cette pratique est interdite par la loi, contrairement à la sortie de l'obstination déraisonnable, qui, elle, a été jugée conforme à l'intérêt du patient. Ainsi, la question du diagnostic différentiel entre les MCS et les UWS devient critique, car c'est bien sur ce mince fil séparant le patient défini comme au moins partiellement conscient et celui dont la récupération apparaît comme impossible que la polémique se crée, source de procédures judiciaires et de débats socio-culturels sur la question de savoir si l'arrêt des traitements constitue ou non une euthanasie. En France, la situation de Vincent Lambert, par exemple, a été qualifiée d'obstination déraisonnable par son médecin, par l'épouse du patient et par plusieurs décisions de justice, tandis que les parents de monsieur Lambert refusaient l'arrêt des traitements, notamment en raison de leurs convictions religieuses, entraînant débats et actions en justices pendant près de 11 ans avant le décès du patient par arrêt des soins le 11 juillet 2019.

De nombreux outils d'évaluation de l'état de conscience ont été développés pour tenter de poser le diagnostic (Tableau 1.4), mais le plus répandu d'entre eux est la *Glasgow Coma Scale*, pratiqué dès l'admission du patient et durant son hospitalisation afin d'estimer son évolution [50]. Toutefois, elle est très rapidement la cible de critiques, car si elle comprend différentes parties (visuelle, motrice et verbale), le décompte final est obtenu par l'addition des différents scores, lissant l'hétérogénéité observée d'un patient à l'autre pour un même score [51], et masquant le poids relatif attribué à ces sous-parties [52]. En outre, il n'existe aucun manuel d'utilisation standardisant l'application de cet examen, faisant reposer son efficacité en grande partie sur l'expertise et l'expérience du praticien.

Des progrès ont été faits par la suite avec notamment la *Coma Recovery Scale-Revised*, outil standardisé utilisé pour établir une distinction UWS vs MCS par l'évaluation de l'éveil, de la perception auditive et visuelle, des capacités motrices et de la capacité à communiquer, en les plaçant sur une échelle allant de la simple activité réflexe à l'activité cérébrale de haut niveau [53, 54]. Malgré ces avancées, il a été démontré en 2009 qu'environ 40% des patients classés en UWS présentaient en réalité des signes de conscience [55].

Ainsi, de nouvelles méthodes de diagnostic différentiel apparaissent comme indispensables dans ces pathologies neurologiques et psychiatriques, pour une prise en charge thérapeutique plus rapide et plus adaptée, traitant des différents enjeux médicaux, émotionnels et sociaux.

L'EEG AU SERVICE DU DIAGNOSTIC DIFFÉRENTIEL

Dans cette thèse, nous avons utilisé la technique de l'ÉlectroEncéphaloGraphie. L'EEG constitue un outil d'évaluation diagnostique des modifications cérébrales pathologiques, permettant de détecter des anomalies du fonctionnement cérébral. Mais il constitue également une méthode de recherche dans le domaine des neurosciences cognitives.

Malgré l'apparition de nombreuses techniques modernes de neuro-imagerie, les avantages de l'EEG continuent d'en faire une des méthodes les plus utilisées pour étudier l'architecture des processus cognitifs. En particulier, grâce à son excellente résolution temporelle, de l'ordre de la milliseconde, l'EEG permet de suivre le décours temporel des processus mentaux, et d'étudier la dynamique des phénomènes cérébraux. De plus, cette technique est entièrement non invasive, peu coûteuse, et transportable au lit du patient.

2.1/ HISTORIQUE : DU GALVANOMÈTRE À L'EEG HAUTE RÉOLUTION

Richard Caton (1842-1926) fut le premier en 1875 à enregistrer l'activité électrique du cortex, attachant des électrodes aux hémisphères d'un cerveau d'animal, et montrant que le courant électrique pouvait changer en fonction du comportement [56]. Par l'usage d'un galvanomètre sensible, il nota l'existence de fluctuations de l'activité cérébrale pendant le sommeil, et l'absence de toute activité suivant la mort.

Bien plus tard, c'est Hans Berger (1873-1941), psychiatre allemand, qui publia en 1929 "*Encephalogram of man*" [57], montrant pour la première fois de l'EEG tel qu'on le connaît aujourd'hui. Il réalisa, le 6 juillet 1924, pendant l'opération neurochirurgicale d'un adolescent de 17 ans, le premier enregistrement de l'activité électrique cérébrale sur un être humain, directement sur la surface du cortex : il s'agissait d'un ÉlectroCorticoGramme. Il parlera ensuite en 1929 des fluctuations de cette activité sous les termes d'ondes alpha et bêta, des oscillations neuronales montrant que des milliers de neurones étaient synchronisés et pouvaient décharger ensemble et qu'ils pouvaient générer des champs électriques détectables et mesurables par EEG. Dès 1939, les termes d'ondes alpha (au début connues sous le nom de "rythme de Berger") et bêta commencèrent à être utilisés pour décrire le taux d'activité neuronale et discuter des applications dans la recherche sur l'épilepsie. Devant les possibilités ouvertes par l'EEG dans l'étude clinique de l'épilepsie

et du sommeil, le premier laboratoire de recherche en EEG clinique fut fondé aux Etats-Unis dans les années 1940, tandis qu'en 1947, l'*American EEG Society*, devenue plus tard *American Clinical Neurophysiology Society* fut créée.

À partir des années 1950, la discussion se centra sur le développement du matériel d'enregistrement (électrodes, amplificateurs, *etc...*) et informatique (processeurs, mémoire, *etc...*) vers une automatisation du traitement des signaux et une meilleure détection des ondes. Une des plus grandes innovations fut le passage d'un enregistrement analogique à digital, s'effectuant non plus à l'aide d'un rouleau de papier, mais par le moyen d'ordinateurs. À présent, des logiciels procurent un environnement tout-en-un pour présenter les stimuli, enregistrer l'activité électrique et effectuer le traitement des données ainsi acquises. Les progrès actuels en EEG ont été possibles grâce à la diminution du coût du calcul par ordinateur et du coût de stockage des données, permettant à la recherche en EEG d'accéder à des projets typiques du "*big data*", avec des capacités analytiques plus rapides. Tout cela mis ensemble permet à la recherche en EEG d'avoir une base robuste et rigoureuse pour la vague suivante de découvertes en psychophysique et en neurosciences.

L'analyse temporelle a prédominé pendant les deux premières décennies après Berger, jusqu'à deux avancées significatives. Tout d'abord, l'utilisation des "grandes moyennes", autrement nommé "moyennage d'ensemble", *i.e.* la moyenne effectuée sur plusieurs présentations du stimulus, a permis de mieux visualiser les potentiels évoqués [58]. Puis sont arrivés les *personal computer*, notamment l'implémentation de la *Fast Fourier Transform* (FFT) permettant une représentation rapide et fiable des fréquences EEG.

Par la suite, sont arrivées des études sur les motifs temps-fréquence et l'introduction des ondelettes dans les analyses EEG [59, 60]. Des systèmes modernes d'acquisition de données, et de nouveaux modèles d'électrodes ont permis l'enregistrement à plus grande échelle, avec davantage de précision ; tandis que des ordinateurs de plus en plus puissants permettent des analyses plus complexes, donnant naissance au champ des neurosciences computationnelles.

Certains pionniers s'intéressèrent alors à l'analyse spatiale. Des filets EEG avec une grande densité d'électrodes sont apparus avec [61–64], dans le but, au départ, d'analyser la topographie des ondes alpha. De hautes résolutions temporelles et spectrales sont atteintes via des études cérébrales de la transformée de Hilbert [65].

À la lecture de l'historique de ce domaine, il est aisé de constater que les percées en traitement du signal EEG ont été effectuées principalement par des chercheurs issus du milieu mathématique, physique ou informatique appliquant leurs connaissances aux neurosciences. Toutefois, il est difficile de maîtriser ces différentes disciplines, chacune ayant sa propre façon d'approcher un problème : les biologistes ont tendance à valoriser l'observation et les expérimentations, tandis que les physiciens raisonnent en termes de modèles théoriques comparés aux résultats expérimentaux, et que les mathématiciens privilégient le raisonnement logique, applicable à n'importe quel objet. Travailler dans le domaine de l'information neurale implique de combiner ces approches et de discuter avec des chercheurs issus d'horizons différents. Pour ce faire, il est nécessaire de connaître les bases de ces domaines voisins et de comprendre leur façon de réfléchir.

2.2/ GÉNÉRATEURS DE L'ACTIVITÉ ÉLECTRIQUE CAPTÉE PAR EEG

Le cerveau humain est constitué d'environ 100 milliards de cellules nerveuses, nommées neurones, qui maintiennent une charge électrique dans le cerveau. Les neurones partagent de nombreuses caractéristiques avec les autres types de cellules, mais leur caractère électrochimique leur permet de transmettre des signaux électriques et de se transmettre des messages sur de longues distances.



FIGURE 2.1 – Vue de profil d'un casque EEG haute résolution à 256 électrodes (HydroCel Geodesic Sensor Net, EGI).

L'ÉlectroEncéphaloGraphie est l'enregistrement de cette activité électrique cérébrale, mesurée comme la différence de sommation de l'activité électrique entre une électrode de référence et une électrode d'enregistrement située en un autre point du scalp (Figure 2.1) ; tandis que l'ÉlectroEncéphaloGramme représente la transcription, sous forme d'un tracé, des variations dans le temps des potentiels électriques recueillis sur la boîte crânienne en différents points du scalp. L'activation des neurones provoque donc des phénomènes électriques qui peuvent être enregistrés à distance. Cependant, l'EEG ne permet de révéler qu'une petite partie de l'activité cérébrale électrique. Il est donc indispensable de comprendre quelles activités l'EEG est susceptible d'examiner ou non.

L'EEG, ainsi que ses aspects liés aux événements, est essentiellement un reflet des courants extracellulaires de masse des réseaux neuronaux d'un large éventail de systèmes cérébraux, et fournit ainsi une fenêtre non invasive particulièrement directe et intégrative sur le fonctionnement du cerveau humain. Au cours du siècle qui a suivi la découverte de l'EEG humain, notre compréhension neurophysiologique de l'activité électrique du cerveau a progressé au niveau microscopique et macroscopique et a été liée à des principes physiques, comme le résumant les manuels standards [66–69].

La présente section s'appuie sur ces textes pour décrire les générateurs d'EEG, dont beaucoup sont applicables à des activités spontanées ou liées à des événements [70–72]. En particulier, il est essentiel, dans le cadre de la neuroimagerie électrique, de savoir

quels événements neuronaux sont détectables, à quel moment et à quelles échelles spatiales.

2.2.1/ DIPÔLES DE COURANT ET ORGANISATION SPATIALE DES NEURONES

Les principaux générateurs de l'EEG sont des plaques étendues de matière grise, polarisées par une entrée synaptique synchrone, soit de manière oscillatoire, soit sous forme d'activité évoquée transitoire, qui reflète des générateurs supplémentaires plutôt qu'une simple réorganisation de phase des oscillations en cours [73].

Dans le cortex, de telles plaques contiennent des milliers de colonnes corticales, où de grands neurones pyramidaux sont alignés perpendiculairement à la surface corticale, et où les différentes couches sont caractérisées par des connexions synaptiques provenant de structures cérébrales différentes. Les courants extracellulaires qui circulent entre les couches de polarité différente sont accompagnés de courants intracellulaires inversés. Les modèles qui tiennent compte des géométries cellulaires réalistes suggèrent que non seulement les dendrites apicales mais aussi basales peuvent contribuer à l'EEG [74].

Les courants résultants sont modélisés sous forme de dipôles de courant dont la direction est donnée par l'orientation des dendrites en question : un dipôle radial (perpendiculaire à la surface crânienne) reflétera les activations des *gyri*, tandis qu'un dipôle tangentiel correspondra aux activations électriques des sillons (Figure 2.2). Ainsi, les champs électriques produits par tous les neurones vont s'organiser en structures ; si ces neurones sont en diverses orientations, les dipôles unitaire vont s'annuler et le champ électrique résultant sera nul, tandis que s'ils sont orientés parallèlement, une sommation de l'activité se produira qui rendra alors celle-ci enregistrable [76]. Par exemple, les structures plus internes telles que les noyaux gris centraux, plus éloignées du cuir chevelu et composées de cellules disposées en étoile induisent des courants qui ne peuvent pas se propager jusqu'aux capteurs situés en surface.

Potentiels négatifs ou positifs

L'orientation de l'aire cérébrale génératrice d'activité électrique par rapport à la surface du cuir chevelu a une influence sur l'orientation de la déflexion mesurée par l'EEG. En effet, un dipôle de courant va être orienté dans un sens ou un autre, donc une électrode à un endroit donné peut enregistrer une activité provenant soit du pôle positif du dipôle, soit du pôle négatif. Ainsi, le caractère positif ou négatif de la déflexion enregistrée en surface par l'électrode ne permet pas de déterminer si une onde correspond à une activité excitatrice ou inhibitrice, ou le niveau de profondeur de la source de courant dans le cortex cérébral. Toutefois, l'enregistrement au niveau d'une même électrode d'une négativité ou d'une positivité correspond bien à deux activités électrophysiologiques distinctes.

Activités électriques à l'origine de l'EEG

— Potentiels d'action (PA)

Le PA est un événement qui survient lorsque des échanges d'ions modifient rapidement les gradients électrochimiques de l'axone neuronal [77], permettant à un signal électrique de se propager rapidement le long de celui-ci à une vitesse allant de quelques mètres à quelques dizaines de mètres par seconde [78], pour une

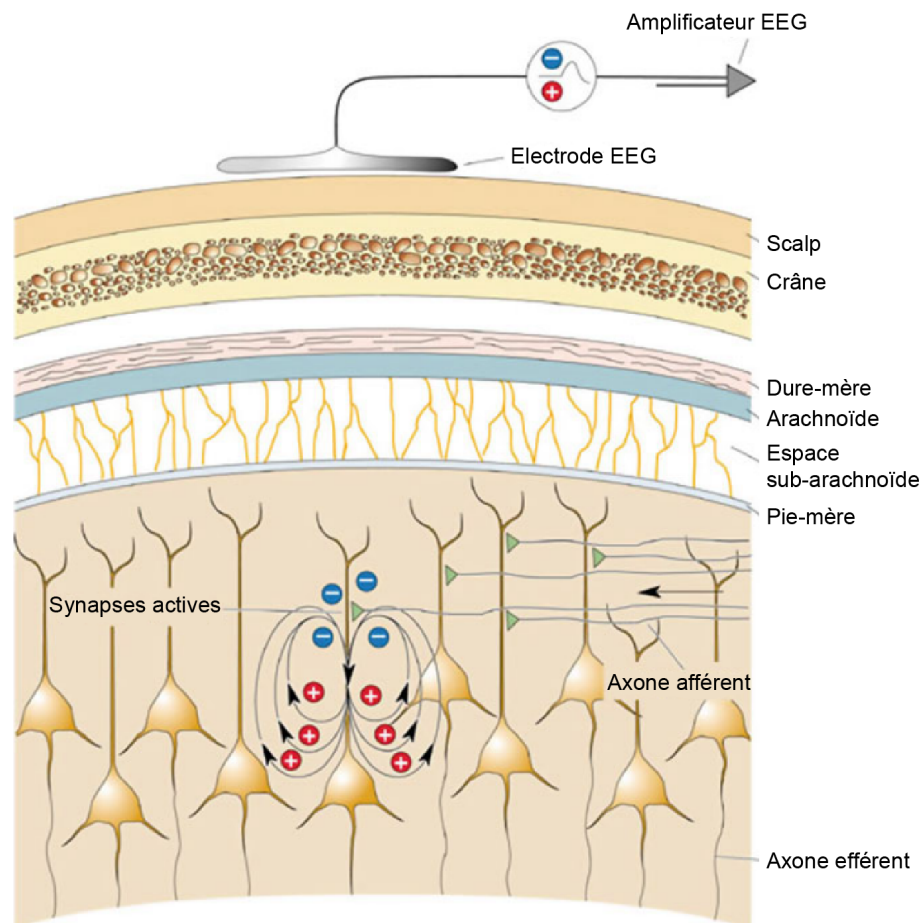


FIGURE 2.2 – Illustration de la génération de très petits champs électriques par des courants synaptiques dans des cellules pyramidales. L'électrode EEG mesure le signal à travers les épaisses couches de tissus. Ce n'est que si des milliers de cellules présentent simultanément leurs petites tensions que les signaux deviennent suffisamment importants pour être visibles à la surface du cuir chevelu. Traduit de Siuly, Li, et Zhang (2016) [75].

durée de 1 à 3 ms et une amplitude de 100 mV environ. Bien qu'ils soient très connus, ces potentiels électriques ne contribuent que très peu à l'activité EEG, car ils sont très courts dans le temps.

— Potentiels Post-Synaptiques (PPS)

Les PPS sont quant à eux générés au niveau des synapses, et s'atténuent avec la distance, pour une amplitude initiale d'environ 10 mV. Bien que beaucoup plus faibles que les PA, les PPS ont une durée bien plus longue (de l'ordre de 20 ms), qui leur permet de contribuer à l'activité EEG de manière bien plus importante que les PA [67].

Il existe ainsi des PPS inhibiteurs (ouverture des canaux K^+ et surpolarisation de la membrane plasmique du neurone), excitateurs (ouverture des canaux Na^+ et dépolarisation de la membrane). Toutefois, les PPS ne sont pas toujours mesurables. Tout d'abord, les canaux Cl^- atténuent l'effet de dépolarisation induit par l'ouverture des autres canaux, donnant lieu à une activité électrique à laquelle l'EEG est aveugle, et d'autre part, la génération de courants extracellulaires repérables

par l'EEG n'a lieu que dans le cas des synapses à gradient électrochimique : les synapses électriques ne générant exclusivement que des courants intracellulaires, leur activité est silencieuse à l'EEG.

Un grand nombre de neurones est nécessaire

L'unité de base générant le signal électrophysiologique cérébral est le neurone. Toutefois, le courant induit par cet unique neurone reste indétectable à l'EEG. Seules de grandes populations de neurones actifs en même temps peuvent générer une intensité électrique cumulée susceptible d'être détectée sur le scalp [79–82].

2.2.2/ POTENTIELS ÉVOQUÉS

La perception d'un stimulus induit un potentiel dit exogène. Si la perception entraîne une activité cognitive, ce potentiel peut être suivi d'un ou plusieurs autres potentiels plus tardifs. Ce sont les potentiels évoqués (*Event Related Potentials*, ou ERPs). L'application de l'EEG aux neurosciences cognitives a nécessité de faire ressortir de l'activité spontanée du cerveau liée à la perception d'un stimulus, et au traitement d'informations requis pour accomplir une tâche. En sommant des tracés issus de la répétition de la même tâche, il est possible de faire apparaître des composantes spécifiques, les réponses évoquées [58]. Ces ERPs sont verrouillés en phase (*phase locked*) avec les stimuli, et leur latence peut donc être déterminée précisément.

Nous proposons donc à présent un tour d'horizon des potentiels évoqués cognitifs les plus courants dans la littérature, et un bref état de l'art mettant en lien ces ERPs avec nos études. Nous traiterons de l'usage de ces ERPs dans les pathologies considérées dans les chapitres concernés (voir Partie II Introduction et Partie III Chapitre 6).

N100

C'est une déflexion négative du signal qui apparaît aux alentours de 100 ms en réponse à un stimulus auditif. Cette activation est sensée refléter l'organisation du cortex auditif primaire. D'après Clynes (1969) [83], cette onde survient à la cessation des sons ayant une durée de plus de 100 ms. Cette onde indique des changements dans l'environnement audio et implique ainsi le processus de traitement auditif pré-attentif qui sous-tend la détection du stimulus. Tout comme la N100-on [84], la N100-off est indépendante de la familiarité du stimulus.

MMN

La MMN "*mismatch negativity*" est une onde négative fronto-centrale d'une amplitude comprise en général entre 0.5 et 5 μ V qui apparaît entre 100 et 2050 ms après le stimulus. La MMN est induite de façon tâche-indépendante par des paradigmes de détection des anomalies auditives. Des sons "rares" apparaissant dans une séquence de répétition de sons "standards" évoquent cette réponse même si l'attention n'est pas portée à la séquence auditive. Elle représente donc un processus cérébral automatique [85]. Les sources cérébrales principales de la MMN sont dans les cortex auditifs temporeux [86]. Toutefois il est aussi décrit une implication du lobe frontal, possiblement due à

un détournement involontaire de l'attention quand le son rare apparaît. Il existe également des générateurs de source dans le thalamus et l'hippocampe [86].

P300

La P300 est une onde positive apparaissant entre 300 et 600 ms après la présentation d'un stimulus rare et constituée de 2 sous-composantes P3a et P3b.

La P3a peut être enregistrée après la présentation de stimuli distracteurs, avec une amplitude maximale dans les aires centro-pariétales. Elle est supposée refléter une allocation de l'attention automatique vers un stimulus nouveau [87]. La P300 proprement dite, la P3b, est reliée à l'allocation volontaire de l'attention envers une cible. La latence de la P3b est souvent considérée comme reflétant le temps d'évaluation du stimulus présenté [88], alors que l'amplitude est liée à la quantité de ressources allouées [89].

N400

Découverte par Kutas et Hillyard [90], la N400 est un potentiel négatif dont le pic se présente environ 400 ms après la présentation du stimulus, observable entre environ 250 et 550 ms. Dans l'article d'origine, la N400 a été mise en évidence par des violations sémantiques, conduisant à la fausse impression qu'il s'agissait d'un détecteur d'anomalies. Toutefois, la N400 est également observée pour des stimuli non linguistiques mais qui font sens : les sons de l'environnement tels que des cris d'animaux ou des sonneries de téléphone [91, 92], des dessins ou des peintures, des expressions faciales, des films, et des gestes. En général, les méthodes utilisées pour induire une N400 fiable incluent des paradigmes d'amorçage ou des attentes fortes. Différents travaux ont démontré que le sujet n'a pas besoin de diriger son attention sur la stimulation [93, 94].

2.3/ AVANTAGES ET INCONVÉNIENTS D'UN PROTOCOLE DE RECHERCHE CLINIQUE BASÉ SUR L'EEG

2.3.1/ PALLIER DES CONTRAINTES DE NEUROIMAGERIE AVEC L'EEG

La mise en évidence de la corrélation d'une tâche mentale avec une variable physiologique cérébrale est complexe. En effet, même s'il existe une activité électrique caractéristique, celle-ci reste enfouie parmi de nombreux signaux non spécifiques dus à l'activité autre du cerveau, qui constitue alors une interférence.

En outre, l'activité mentale, même spécifique, donne des signaux non stationnaires car elle dépend de l'état de l'individu au moment de la mesure : concentration, fatigue, émotions ressenties, pensées sans rapport avec la tâche expérimentale *etc.* Les tâches cognitives réalisées par le sujet doivent donc être sélectionnées de façon à induire une activité spécifique suffisamment forte pour être enregistrable et différenciable du "bruit de fond".

Toutefois, définir une tâche mentale précise ne suffit pas pour obtenir un neuromarqueur robuste. Une variation du signal EEG doit être localisée dans le temps, autrement dit corrélée au moment où l'état mental se modifie, et dans l'espace via des électrodes et/ou une localisation des structures cérébrales sources de courant spécifiques).

Les caractéristiques du système de mesure EEG-HR sont essentielles pour acquérir un signal susceptible de retranscrire l'activité cérébrale d'intérêt :

- résolution spatiale : localisation de la source de production de l'activité, dont la précision est directement corrélée au nombre d'électrodes
- résolution temporelle : afin d'obtenir une corrélation précise entre un état mental enregistré et le stimulus ou la réponse ayant supposément induit cet état mental.

Nos connaissances sur la fonction cérébrale ont été grandement améliorées, dans les dernières décennies, par le développement de différentes techniques d'enregistrement, du niveau microscopique, avec l'étude de l'activité synaptique, au niveau macroscopique avec des méthodes de neuroimagerie telles que la MagnétoEncéphaloGraphie (MEG), la Tomographie par Émission de Positons (PET) et l'IRM fonctionnelle (IRMf) (Figure 2.3). Pourquoi, dès lors, continuer à se concentrer sur le prédécesseur de ces méthodes ?

À l'heure actuelle, l'utilisation de l'EEG s'est répandue grâce à ses avantages intrinsèques : comme méthode de neuroimagerie relativement peu chère et transportable au lit du patient.

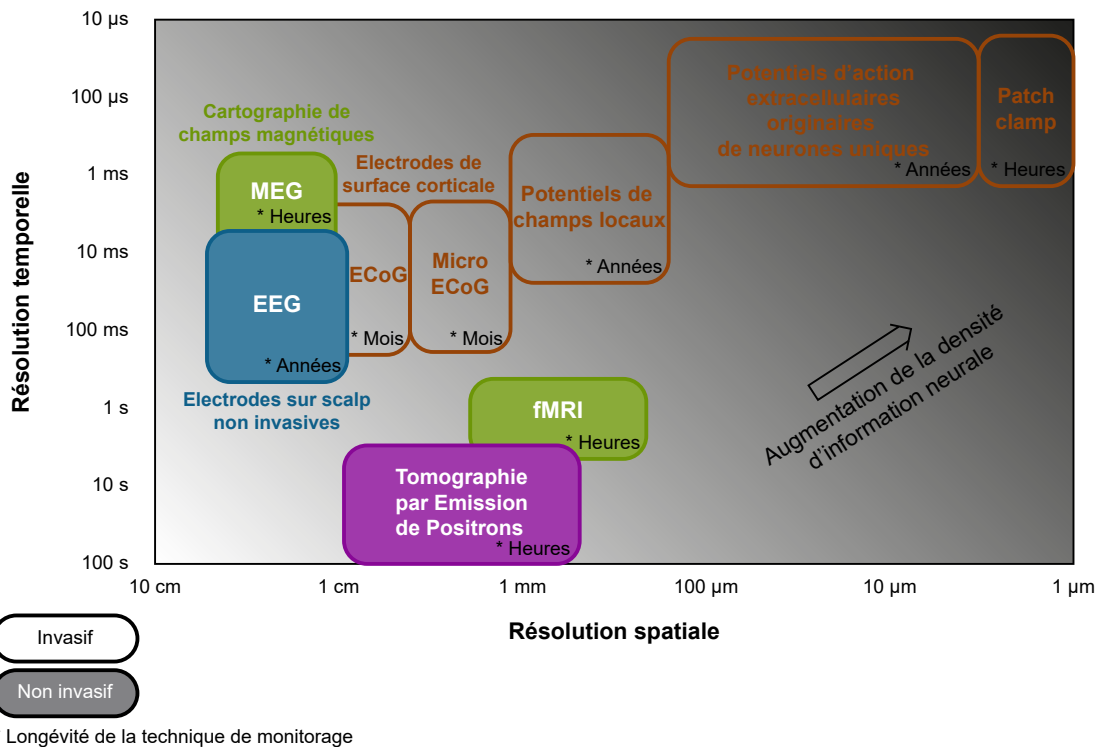


FIGURE 2.3 – Illustration de la résolution spatio-temporelle des technologies actuelles de monitoring du cerveau. Adapté et traduit de Muthuswamy, Sridharan et Okandan (2016). [95]

2.3.2/ LIMITES DE L'EEG ET *machine learning* POUR LE DIAGNOSTIC CLINIQUE

Lors de la mesure d'un EEG, nous disposons souvent de grandes quantités de données correspondant à différentes conditions expérimentales, en particulier lorsque les enregistrements sont effectués sur une longue période. Cependant, les procédures actuelles de traitement, d'évaluation et de classification de ces signaux ne sont pas satisfaisantes

car de nombreuses techniques utilisées dans la recherche ne sont pas suffisamment standardisées pour être utilisées dans le contexte clinique. En effet, l'évaluation d'un enregistrement EEG est généralement effectuée par des spécialistes en EEG expérimentés qui analysent visuellement les enregistrements en particulier en EEG de sommeil et en épileptologie [96, 97]. Il n'existe pas de critères standards pour évaluer par exemple quels essais peuvent être rejetés pour cause d'artéfacts ou conservés pour analyse. Ce fonctionnement entraîne plusieurs conséquences :

- l'acquisition de l'expérience et des connaissances nécessaires au traitement des données par un spécialiste est coûteuse en temps et en ressources,
- le processus de traitement en lui-même est également très long, impliquant de mobiliser un spécialiste sur un seul projet pendant un certain temps,
- la longueur du traitement peut en outre induire chez le spécialiste une fatigue ou une habitude aux données, sources d'erreurs,
- les résultats du traitement peuvent varier d'un spécialiste à un autre, diminuant la reproductibilité de l'étude,
- les résultats du traitement peuvent varier pour un même spécialiste entre deux analyses sur les mêmes données.

Pour extraire des informations d'une telle quantité de données, des systèmes d'analyse automatisés sont nécessaires pour analyser et classer les données à l'aide de techniques appropriées. Il est donc essentiel de développer des systèmes automatiques de classification des signaux EEG enregistrés afin de soutenir le diagnostic des maladies du cerveau et de contribuer à une meilleure compréhension des états mentaux pour les applications cliniques. Dans le chapitre suivant, nous verrons comment les méthodes d'apprentissage machine, dites "d'intelligence artificielle" peuvent trouver leur place dans les processus de traitement du signal, d'identification des artéfacts et de classification des états cérébraux. Ces méthodes rejoindront à terme les outils de diagnostic en clinique.

UNE NOUVELLE APPROCHE DE L'ANALYSE EEG : LE *machine learning*

Dans ce travail de thèse, l'objectif est d'appliquer ces nouvelles manières de traiter le signal à des signaux EEG. Ces méthodes ne sont pas encore standardisées car elles sont relativement récentes, aussi, ce chapitre a pour objet de détailler les étapes clés de l'application du *machine learning* et les spécifications particulières concernant les signaux EEG.

3.1/ DÉFINITIONS

3.1.1/ L'ORDINATEUR ET L'ALGORITHME

“Un ordinateur est une machine, munie d'une mémoire dans laquelle on insère des données (ou informations, d'où le nom informatique) et un algorithme, codé dans un langage propre, qui définit le traitement à effectuer sur ces données. La machine ne peut décider elle-même d'exécuter telle ou telle consigne, elle ne peut qu'exécuter les étapes déterminées à l'avance par l'auteur humain de l'algorithme.” [98].

3.1.2/ L'INTELLIGENCE ARTIFICIELLE (IA)

D'après Marvin Minsky, l'intelligence artificielle est : “la discipline de l'informatique qui s'intéresse à la construction de programmes informatiques qui s'adonnent à des tâches qui sont, pour l'instant, accomplies de façon plus satisfaisantes par des êtres humains, car elles demandent des processus mentaux de haut niveau tels que : l'apprentissage perceptuel, l'organisation de la mémoire et le raisonnement critique.”

L'IA consiste donc à écrire des algorithmes plus sophistiqués, qui calculent automatiquement les bonnes réponses à partir des données fournies, au lieu d'écrire soi-même tout le détail des instructions à la main. Ce principe est au cœur de la technique dite “d'apprentissage machine”, ou “*machine learning*”. Le traitement des données est décrit dans un programme qui est lui-même une donnée fournie à la machine. Ce programme peut ainsi être modifié ou construit par un autre programme. Le résultat fourni n'est donc pas créé de toutes pièces par l'IA, mais dépend des instructions données par l'informaticien [98].

Un ordinateur ne peut écrire lui-même un programme pour résoudre un problème pour lequel il n'a pas été conçu car il y a trop de données différentes, trop de règles requises pour que ce programme fonctionne correctement. Le programme d'IA qui résoudrait tous les problèmes n'existe pas car, justement, chaque problème nécessite un programme spécifique et des données adaptées [98].

L'application du machine learning à des signaux EEG nécessite la réalisation d'un certain nombre d'étapes :

3.2/ ÉTAPE 1 : DÉFINITION DU PROBLÈME À RÉSOUDRE

Dans nos deux études, nous avons cherché à classer les enregistrements EEG en fonction de la familiarité des chansons écoutées (familières ou inconnues), d'une part ; et la mémoire sémantique liée au contexte (congruent ou incongruent) d'autre part. Il nous a fallu choisir un modèle d'apprentissage permettant cette classification. Dans un premier temps, nous proposons de définir les grandes catégories de *machine learning* afin de déterminer celle qui sera la plus adaptée à l'atteinte de notre objectif.

3.2.1/ TYPES D'APPRENTISSAGES

3.2.1.1/ APPRENTISSAGE NON SUPERVISÉ

L'apprentissage machine non supervisé permet de découvrir des modèles cachés ou des structures intrinsèques dans les données. Il est utilisé pour tirer des conclusions à partir d'ensembles de données constitués de données d'entrée sans réponses étiquetées [99]. Le *clustering* est la technique d'apprentissage non supervisé la plus courante. Elle est utilisée pour l'analyse exploratoire des données afin de trouver des modèles cachés ou des regroupements dans les données. Les applications du *clustering* incluent par exemple l'analyse des séquences de gènes, les études de marché, la reconnaissance d'objets, *etc...* Dans le cas des données EEG, nous connaissons à l'avance la catégorie à laquelle appartient chacun des essais enregistrés, puisque nous induisons les différents états cérébraux par des stimuli externes à l'individu. L'apprentissage non supervisé n'est donc pas *a priori* le plus approprié pour ce type d'études.

3.2.1.2/ APPRENTISSAGE SUPERVISÉ

L'objectif de l'apprentissage machine supervisé est de construire un modèle qui fait des prédictions basées sur des preuves en présence d'incertitude. Un algorithme d'apprentissage supervisé prend un ensemble connu de données d'entrée et de réponses connues à ces données (étiquettes) et forme un modèle pour générer des prédictions raisonnables pour la réponse à de nouvelles données [100, 101]. L'étiquetage va indiquer à la machine si la valeur à prédire pour une observation précise est telle ou telle catégorie ou condition expérimentale, ou une valeur numérique.

Régression

Les techniques de régression permettent de prévoir des réponses numériques continues, par exemple des changements de température ou des fluctuations de la demande d'énergie. Les applications typiques comprennent entre autres la prévision de la charge électrique et le commerce algorithmique dans la finance.

Classification

Les techniques de classification permettent de prédire des réponses discrètes, par exemple si une tumeur est cancéreuse ou bénigne. Autrement dit, les modèles de classification classent les données d'entrée dans des catégories prédéterminées. Ce type de modélisation est retrouvé par exemple en imagerie médicale, en reconnaissance vocale *etc.* En neuroimagerie, l'objectif de la classification est d'attribuer les vecteurs de caractéristiques à l'une des n catégories discrètes, où n correspond au nombre d'états cérébraux différents que le système est capable de détecter.

L'algorithme de classification peut être mis en œuvre pour distinguer deux classes (problèmes binaires), ou plusieurs classes (problèmes multi classes). En ce sens, étant donné que la difficulté de la tâche de classification augmente avec le nombre de classes, il y a un compromis à faire entre fournir un outil plus complet mais au résultat plus hasardeux, ou assurer une performance efficace avec une portée limitée. Étant donné que nos études comprennent la plupart du temps deux conditions expérimentales, un algorithme de classification binaire est tout indiqué.

3.2.2/ ALGORITHMES DE CLASSIFICATION LES PLUS COURANTS

Il existe un grand nombre d'algorithmes de classification différents, chacun présentant des spécificités le rendant davantage pertinent que ses homologues face à tel ou tel problème. En EEG, les plus fréquemment utilisés sont les suivants :

3.2.2.1/ *K-Nearest Neighbours*

L'algorithme *K-Nearest Neighbours* (les K plus proches voisins) vise à déterminer à quelle catégorie appartient un essai en se basant sur le groupe d'appartenance des K essais les plus proches lorsque l'on projette le jeu d'entraînement dans l'espace des caractéristiques [100, 102]. La probabilité pour un essai de faire partie d'une catégorie est calculée comme le ratio des membres de ladite catégorie parmi les K plus proches voisins [100, 102–104]. La valeur optimale de K est déterminée après les évaluations expérimentales.

3.2.2.2/ *Naïve Bayes*

Le *Naïve Bayes* est un algorithme de classification basé sur la probabilité conditionnelle qu'un événement se produise en fonction d'un événement qui s'est déjà produit. Il est fondé sur le Théorème de Bayes qui s'appuie sur des hypothèses d'indépendance fortes (naïves) : on suppose que tous les essais et sont indépendants les uns des autres pour une classe donnée, et généralement que leurs caractéristiques ont une distribution gaussienne si elles sont continues [75, 105].

3.2.2.3/ Linear Discriminant Analysis

Le *Linear Discriminant Analysis* (LDA) construit un modèle prédictif de l'appartenance d'un essai à une catégorie en utilisant des hyperplans pour séparer les données. L'objectif est de séparer au maximum les groupes et d'éliminer les caractéristiques extraites du signal qui ne sont pas pertinentes pour la discrimination des données. Cette technique nécessite très peu de calculs [100, 105].

3.2.2.4/ Support Vector Machine

Le *Support Vector Machine* (SVM), développé suite aux travaux de Vladimir Vapnik [106], consiste lui aussi à trouver un hyperplan optimal de frontière entre deux classes de données. Cependant, cette méthode est augmentée de la recherche de la marge maximale de chaque côté de cette séparation, optimisant la capacité de généralisation du modèle sur des données inconnues.

La détermination de cet hyperplan optimal passe par la détermination de la distance minimale entre celui-ci et les essais le plus proches dans chacune des deux classes. En d'autres termes, on recherche une frontière de décision telle que les marges soient les plus larges possible. Ces essais qui à eux seuls permettent de définir l'hyperplan sont appelés vecteurs supports (Figure 3.1).

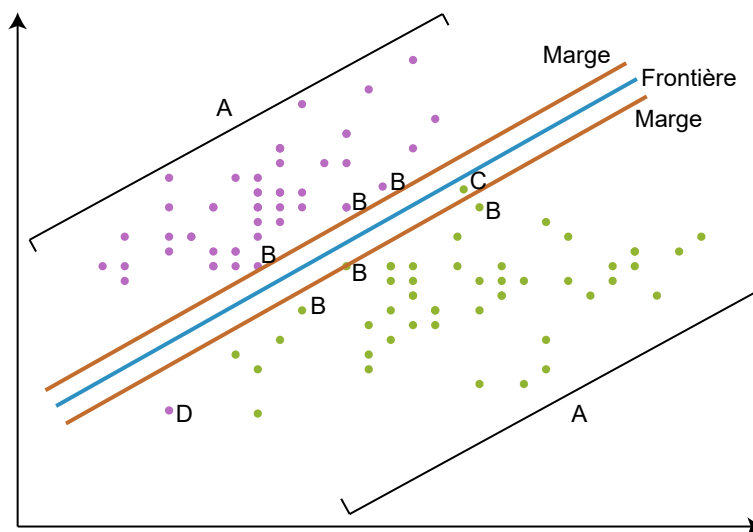


FIGURE 3.1 – Représentation à deux dimensions d'une frontière de séparation construite par le SVM. A. Essai bien classé et pas dans la zone de la marge. Il ne représente pas un vecteur support. B. Essai bien classé et aux frontières de la zone de la marge. Il représente un vecteur support. C. Essai bien classé et dans la zone de la marge. Il ne représente pas un vecteur support. D. Essai du mauvais coté de la frontière, il est mal classé et ne représente pas un vecteur support.

Toutefois, un hyperplan n'est pas toujours la meilleure façon de séparer deux classes de données, par exemple pour des classes réparties à l'intérieur et à l'extérieur d'une hypersphère. La détermination d'une telle fonction non linéaire étant très difficile, voire impossible, les données sont alors projetées dans un espace de dimension plus élevée où cette fonction devient linéaire (Figure 3.2). Cette méthode est applicable par le SVM

par le biais de fonctions dites “de noyau” (*kernel machines*).

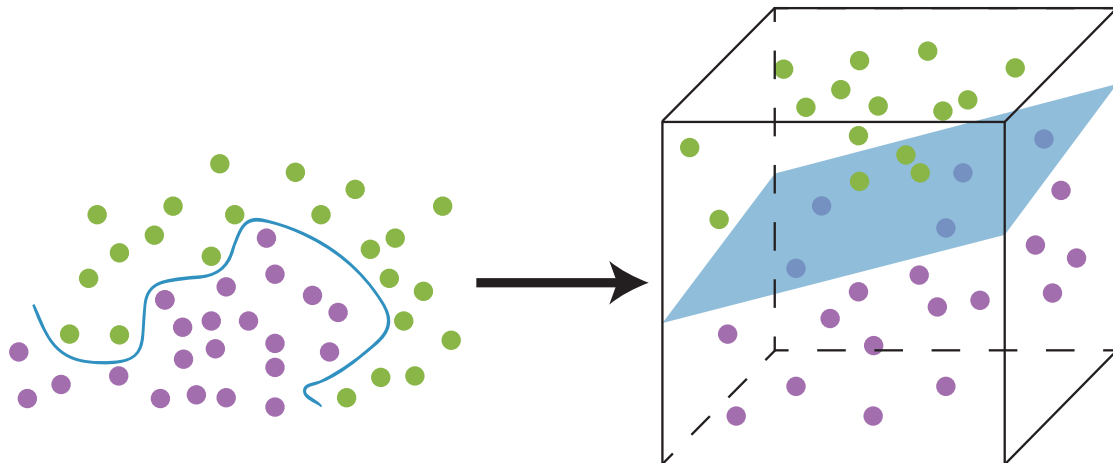


FIGURE 3.2 – Illustration de l’astuce du noyau : les données sont projetées dans un espace de plus grande dimension, où les données peuvent être séparées par un hyperplan. Cette frontière linéaire dans l’espace de grande dimension est équivalente à une frontière non linéaire dans l’espace d’origine.

La construction de tels noyaux est largement étudiée dans [107], néanmoins, il en existe certains qui sont très utilisés et qui sont considérés comme standards (Figure 3.3) :

- *Kernel* linéaire : dans le cas où les données sont linéairement séparables, et qu’un changement de dimension n’est pas requis, la frontière de décision reste un hyperplan.
- *Kernel* polynomial : dans le cas où les données ne sont pas linéairement séparables, et qu’un changement de dimension est nécessaire, la fonction de séparation peut être définie sur la base de fonctions polynomiales.
- *Kernel Radial Basis Function* (RBF) : dans le cas où les données ne sont pas linéairement séparables, et qu’un changement de dimension est nécessaire, la fonction de séparation peut être définie sur la base de fonctions radiales.

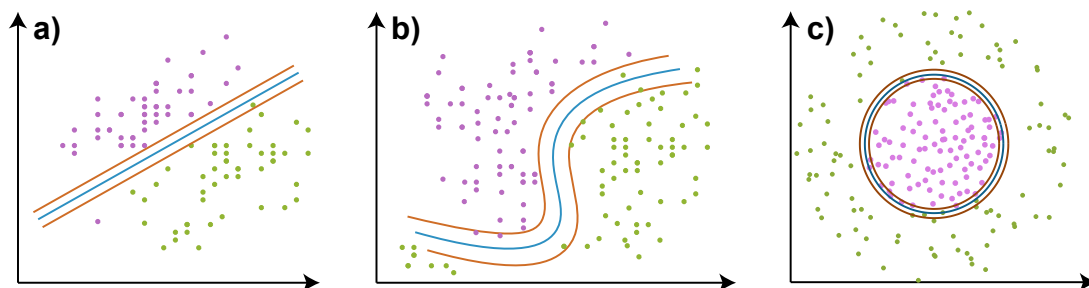


FIGURE 3.3 – Représentation à deux dimensions du type de frontière de séparation possible par l’utilisation de kernels. a) *Kernel* linéaire. b) *Kernel* polynomial. c) *Kernel* RBF.

3.3/ ÉTAPE 2 : ACQUISITION DES DONNÉES D'APPRENTISSAGE ET DE TESTS

3.3.1/ NÉCESSITÉ D'UNE GRAND NOMBRE DE DONNÉES

La capacité du modèle à représenter correctement une population d'essais va dépendre en grande partie de la taille de l'échantillon qui lui sera présenté lors de l'entraînement.

Nous avons vu que pour la classification, nous avons besoin de deux hyperplans de marge à égale distance de l'hyperplan optimal qui sépare les essais. Supposons que nous ayons un ensemble d'essais de dimension d , c'est-à-dire que chaque essai est représenté par un vecteur de d caractéristiques. Puisque nous sommes en dimension d , chaque hyperplan de marge peut être construit en utilisant un minimum de d vecteurs supports.

Ainsi, pour la classification, nous aurons besoin d'un minimum de 2^d vecteurs supports. Plus le nombre de caractéristiques utilisées pour décrire les données sera élevé, plus le nombre d'essais nécessaires pour entraîner le modèle de classification sera important. Par exemple : pour des données décrites par 4 caractéristiques, il nous faudra au minimum 16 essais. Mais plus les données sont complexes, telles que le sont les enregistrements de l'activité cérébrale, plus le nombre de caractéristiques nécessaires à les décrire est grand. Toutefois, les contraintes liées aux examens EEG, notamment la fatigue du sujet, font qu'il est difficile en pratique d'obtenir un très grand nombre d'essais.

3.3.2/ SIMILARITÉ INTRACLASSE ET DISSIMILARITÉ INTERCLASSES

Un autre point important est celui de la similarité intraclasse et de la dissimilarité interclasses. Afin de pouvoir différencier au mieux les essais appartenant à la classe A de ceux de la classe B, il est préférable que les individus regroupés au sein d'une même classe soient le plus semblables possibles (homogénéité intraclasse), tandis que les deux classes doivent être le plus dissemblables possible (hétérogénéité interclasses). Si ces deux conditions ne sont pas respectées, la difficulté d'obtenir une frontière de séparation adéquate et la probabilité de voir des données mal classées augmente.

3.3.3/ ÉQUILIBRE ENTRE LES CLASSES

Il arrive dans certains cas, en particulier dans le domaine du diagnostic, que la quantité de données recueillie pour une classe soit bien inférieure à l'autre, par exemple lorsque l'on dispose de peu de patients par rapport au nombre de sujets sains. Cette situation peut également survenir à la suite du pré-traitement des données EEG, ou en fonction des tâches effectuées par les sujets lors de l'examen EEG, où il arrive alors que le nombre d'essais présents ne soit pas le même dans les deux conditions.

Ce déséquilibre est problématique car les classificateurs sont en général plus sensibles à la détection de la classe majoritaire. Autrement dit, la probabilité de correctement classer un essai au hasard n'est plus de 50%. Considérons l'exemple suivant : si l'on a 1000 essais dans une classe et 10 dans l'autre, la probabilité d'appartenance d'un essai à la classe majoritaire est de 99%, ce qui signifie que le classificateur peut obtenir 99%

3.4. ÉTAPE 3 : NETTOYAGE DES DONNÉES ET EXTRACTION DES CARACTÉRISTIQUES³⁷

de bonnes réponses au hasard, en étiquetant tout simplement tous les essais des deux conditions comme appartenant à la classe majoritaire.

Pour pallier ce défaut, nous avons alors trois possibilités :

- Ignorer le problème : ce qui n'est pas recommandé, mais peut être acceptable lorsque l'écart entre les deux classes n'est pas significatif (exemple : 200 essais dans une classe, 210 dans l'autre).
- Sous-échantillonner la classe majoritaire : en d'autres termes, éliminer une partie des essais de cette condition pour en ramener le nombre à un équivalent de la classe minoritaire. Cette méthode peut toutefois avoir pour inconvénient de limiter la quantité des données disponibles pour entraîner le modèle.
- Sur-échantillonner la classe minoritaire : divers méthodes ont été développées en ce sens, la plus connue étant la *Synthetic Minority Over-Sampling Technique* (SMOTE) [108], qui consiste à créer de nouveaux essais, synthétiques, par interpolations d'essais de la classe minoritaire.

Nous avons testé ces trois méthodes dans nos études, mais les meilleures performances ayant été obtenues dans notre cas par le sous-échantillonnage, nous ne présenterons que les résultats liés à celui-ci.

3.4/ ÉTAPE 3 : NETTOYAGE DES DONNÉES ET EXTRACTION DES CARACTÉRISTIQUES

3.4.1/ NETTOYAGE DES DONNÉES

Les signaux électriques présents dans l'EEG qui sont d'origine non cérébrale sont appelés artefacts. Les données EEG sont presque toujours contaminées par de tels artefacts, dont l'amplitude est relative à celle des signaux d'intérêt. C'est l'une des raisons pour lesquelles il faut une expérience considérable pour interpréter correctement les EEG sur le plan clinique. La figure 3.4 présente une illustration des quatre types d'artefacts les plus courants dans les enregistrements EEG humains.

Il existe, là encore, de très nombreuses techniques d'élimination de ces signaux parasites, parmi lesquels on peut notamment trouver l'inspection visuelle par un expert [109–111], l'analyse en composantes indépendantes (*Independent Component Analysis*, ou ICA) [112–114] ou plus récemment la séparation à l'aveugle des sources (*Blind Source Separation*, ou BSS) [115–117], pour ne citer que les plus connues.

Nous nous sommes intéressés plus particulièrement à une méthode entièrement automatisée, baptisée le débruitage xDAWN. L'objectif du filtre spatial xDAWN est d'améliorer le potentiel lié à l'événement afin qu'il se distingue du bruit, en maximisant le rapport signal/signal plus bruit dans les études à essai unique [118]. Cet algorithme est principalement connu dans le problème de classification BCI dite de "*P300-speller*". La méthode xDAWN non supervisée propose d'estimer d'abord un sous-espace contenant la plupart des potentiels évoqués sur lesquels les époques EEG sont ensuite projetées, ce qui améliore les ERP. Pour les besoins de cette thèse, nous avons écrit l'algorithme nous-même, en suivant les recommandations de Rivet (2009) [118] et en utilisant le logiciel Matlab R2016b[®] de MathWorks[®].

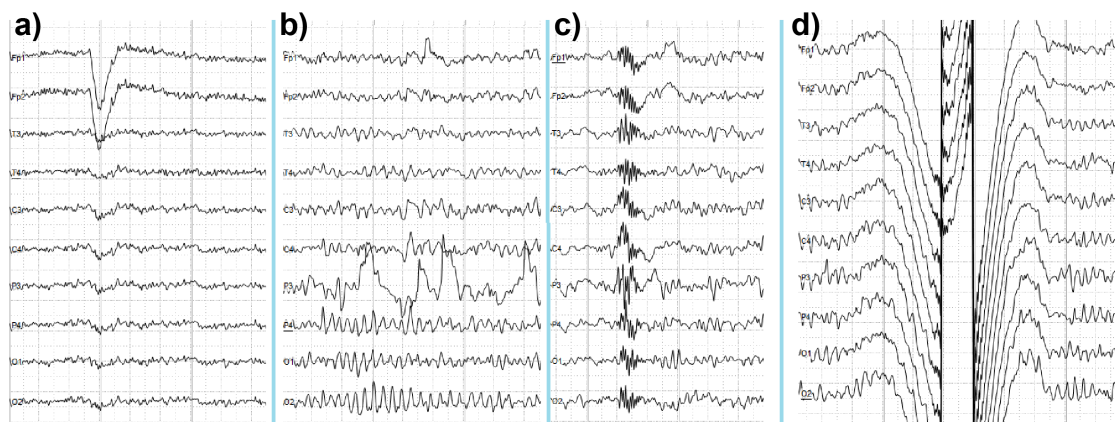


FIGURE 3.4 – Artéfacts les plus couramment rencontrés dans l’analyse du signal EEG. a) Artefact électro-oculographique causé par l’excitation des muscles du globe oculaire (lié au clignement des yeux, par exemple). Grande amplitude, lente, onde positive prédominante dans les électrodes frontales. b) Artefact d’électrode causé par un mauvais contact (et donc une plus grande impédance) entre l’électrode P3 et la peau. c) Artefact de déglutition. d) Artefact de l’électrode de référence commune causé par un mauvais contact entre l’électrode de référence et la peau. Onde énorme similaire dans tous les canaux. Tiré de : <https://en.wikipedia.org/wiki/Electroencephalography>

3.4.2/ EXTRACTION DES CARACTÉRISTIQUES

La phase de préparation des données et d’ajustement du modèle d’apprentissage sont les deux plus importantes. Les signaux EEG recueillis doivent en premier lieu être analysés et préparés afin de leur donner une signification. L’objectif de l’extraction de caractéristiques est de sélectionner des informations pertinentes à partir des signaux enregistrés, qui conféreront un caractère distinctif à l’activité cérébrale qui est déclenchée lors de la tâche correspondante, afin qu’elle puisse être reconnue à partir d’autres essais [119]. Un même jeu de données peut ainsi être exploité de manières différentes.

Dans la plupart des cas, ce processus réduit la dimensionnalité des données en excluant les informations redondantes, et il améliore les attributs du signal en exploitant les aspects qui sont liés aux phénomènes physiologiques. Les valeurs dérivées sont projetées dans l’espace des caractéristiques, où l’on s’attend à ce que les éléments provenant de différentes conditions soient situés de manière aussi séparée que possible, afin que des groupes distincts puissent être facilement identifiés.

Cette phase est particulièrement délicate et fait l’objet d’un grand nombre de recherches afin de déterminer quel serait la méthode idéale. En effet, si les caractéristiques des données fournies à l’algorithme ne sont pas pertinentes, la performance de classification s’en ressentira, selon le principe du “*garbage in, garbage out*”. Malheureusement, aucun consensus ne se dégage vraiment de la littérature, car ce processus est dépendant du type d’information recherché dans le signal, du matériel d’enregistrement EEG utilisé, et, surtout, du protocole d’acquisition des données élaboré au départ. Plusieurs méthodes sont le plus souvent explorées afin de déterminer laquelle donne les meilleurs résultats, et notre étude ne fait pas exception.

3.5. ÉTAPE 4 : ÉVALUATION DE LA PERFORMANCE, AJUSTEMENT DU MODÈLE³⁹

Ainsi que nous le verrons dans la section “Extraction des caractéristiques” du Chapitre 4, la fréquence est l’un des critères les plus utilisés pour évaluer les anomalies des EEG cliniques et pour comprendre les comportements fonctionnels dans la recherche cognitive.

La fréquence fait référence à l’activité rythmique répétitive (en Hz), soit le nombre de cycles par seconde. Chez les adultes en bonne santé, les amplitudes et les fréquences de signaux électriques cérébraux changent d’un état à l’autre, comme l’éveil et le sommeil. Il existe cinq grandes ondes cérébrales qui sont généralement classées en bandes spécifiques telles que 1-4 Hz (delta, δ), 4-8 Hz (thêta, θ), 8-12 Hz (alpha, α), 12-30 Hz (bêta, β) et supérieur à 30 Hz (gamma, γ). Les informations fréquentielles du signal EEG peuvent être extraites, le plus souvent par transformée de Fourier *Fast Fourier Transform*, ou FFT [120], mais aussi par ondelettes [121, 122], par modèles spatiaux communs *Common Spatial Patterns*, ou CSP) [123–125] ou encore par le biais de l’un de leurs nombreux dérivés.

Dans d’autres études, l’utilisation du signal temporel, brut ou non, est privilégiée [126, 127], et de très nombreuses autres méthodes d’extraction de données pertinentes ont été développées ces dernières années. Chacun recherchant la méthode s’appliquant le mieux à son jeu de données, il existe presque autant de sous-espèces de techniques d’extraction des caractéristiques que d’articles.

3.5/ ÉTAPE 4 : ÉVALUATION DE LA PERFORMANCE ET AJUSTEMENT DU MODÈLE

3.5.1/ ÉVALUATION DE LA PERFORMANCE DU MODÈLE APPRIS

Les critères d’évaluation des performances d’une méthodologie sont une partie importante de sa conception. Dans notre étude, la stabilité de la performance d’une méthode de classification est évaluée selon les critères standards, à savoir la précision de la classification, la sensibilité et la spécificité, qui sont définies ainsi (Figure 3.5) [128–131] :

- Précision de la classification ou *accuracy* : la proportion d’essais correctement classés par rapport au nombre total d’essais. Si la précision est de 1 (100%), la capacité de discrimination du modèle est parfaite. Si elle est égale à 0.5, le classificateur n’a aucun pouvoir discriminant et aucun classificateur approprié ne devrait avoir une aire sous cette courbe inférieure à 0,5 [132].
- Sensibilité ou *sensitivity* : le nombre d’essais positifs classés comme positifs divisé par le nombre total d’essais réellement positifs,
- Spécificité ou *specificity* : le nombre d’essais négatifs classés comme négatifs divisé par le nombre total d’essais réellement négatifs.

Il existe un grand nombre d’autres critères d’évaluation de la performance de classification, telles que le F1-score, la g-mean, *etc...* que nous n’employons pas dans ces travaux et que nous ne développerons donc pas ici.

		Catégorie de l'essai	
		A	B
Réponse du classificateur	A	Vrai positif	Faux positif
	B	Faux négatif	Vrai négatif

FIGURE 3.5 – Matrice de confusion permettant de calculer les différentes métriques de performance. Précision = (Vrais positifs + Vrais négatifs) / Nombre total d'essais. Spécificité = Vrais négatifs / (Vrais négatifs + Faux positifs). Sensibilité = Vrais positifs / (Vrais positifs + Faux négatifs).

3.5.2/ VALIDATION DE LA PERFORMANCE

Lors de l'évaluation d'un modèle d'apprentissage machine, il est important de mesurer sa performance de généralisation, c'est-à-dire de savoir s'il est capable d'obtenir une bonne performance sur des données qui lui sont inconnues. Pour ce faire, il est courant de diviser les données disponibles en un ensemble destiné à l'entraînement du modèle, et en un autre destiné à le tester. Lorsque les hyperparamètres doivent être réglés, la performance de l'ensemble de test ne peut plus être utilisée comme une évaluation impartiale des performances de généralisation du modèle. Par conséquent, l'ensemble de formation est divisé pour obtenir un troisième ensemble appelé "ensemble de validation" qui est utilisé pour sélectionner le meilleur hyperparamètre, en laissant l'ensemble de test de côté pendant ce processus afin d'évaluer les performances du meilleur modèle de manière impartiale. Toutefois, lorsque la quantité de données disponibles est faible, la division des données en différents ensembles et le fait de n'utiliser qu'un sous-ensemble pour l'entraînement peuvent sérieusement nuire à la performance des modèles gourmands en données. Une procédure appelée "validation croisée" est utilisée dans ces cas, où les données sont ventilées en différentes partitions, qui seront ensuite utilisées successivement comme données d'entraînement ou de validation (Figure 3.6).

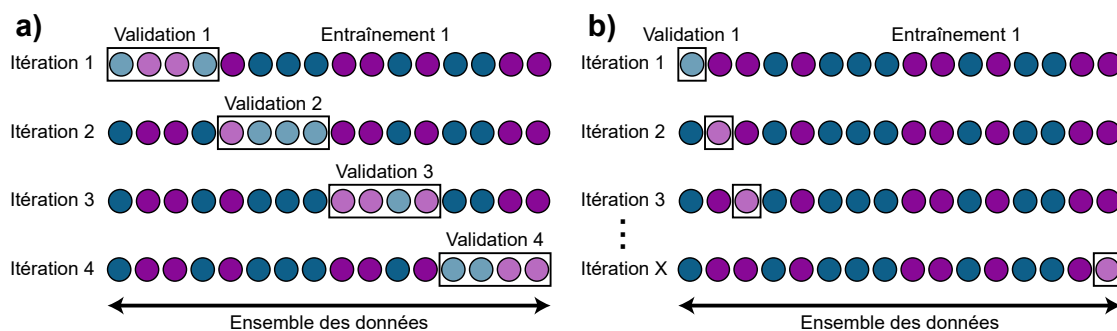


FIGURE 3.6 – Architectures de validation croisée les plus utilisées. a) Validation croisée *k-fold* : le jeu d'entraînement est divisé en sous-ensemble, dont chacun est utilisé à son tour pour valider le modèle produit par l'entraînement du modèle SVM sur tous les autres sous-ensembles. b) Validation croisée *Leave-One-Out*, ou *Leave-One-Subject-Out* : cette fois le sous-ensemble de validation du modèle ne comporte qu'un seul essai ou sujet.

3.5.3/ AJUSTEMENT DU MODÈLE D'APPRENTISSAGE

Selon le *kernel* utilisé, on peut devoir ajuster un ou plusieurs hyperparamètres régularisant l'activité de celui-ci. Le premier est commun aux *kernels* linéaire et RBF : le paramètre C permet de faire un compromis entre la classification correcte des exemples de formation et la maximisation de la marge de la frontière de décision. Pour des valeurs plus élevées de C, une marge plus faible sera acceptée, avec une frontière de séparation plus complexe et qui permettra de classer correctement un maximum d'essais à l'entraînement, tandis qu'un C plus faible encouragera une marge plus importante, donc une fonction de décision plus simple, au détriment de l'*accuracy* d'entraînement. Autrement dit, C est un paramètre qui permet de faire un compromis entre la classification erronée des essais d'entraînement et la simplicité du modèle ainsi créé.

Dans le cas du *kernel* RBF, un second hyperparamètre entre en jeu : sigma. En termes simples, sigma est le rayon de la zone d'influence des vecteurs de support sélectionnés : si sigma est très faible, il ne pourra encapsuler que le vecteur support lui-même, tandis qu'en revanche, si sigma est très grand, le vecteur de support influencera les essais plus éloignés.

Le choix de ces hyperparamètres va donc fortement influencer la performance du modèle, et il est extrêmement recommandé de rechercher ceux qui seront le plus adaptés, sous peine d'obtenir un modèle sous- ou sur-ajusté aux données d'entraînement, et ainsi incapable de produire des résultats convenables sur les données de test. En outre, puisque les valeurs optimales des hyperparamètres seront différentes selon le jeu de données d'apprentissage, il est nécessaire de les ajuster séparément pour chaque tâche. Le plus souvent, on emploie une méthode de passage au crible (*gridsearch*) entre deux bornes fixées à l'avance pour chacun de ces hyperparamètres, où l'on examine les résultats et le comportement du modèle en fonction des valeurs d'hyperparamètres retenues [133–135].

Quelques particularités sont à noter :

- Dans le cas du RBF, différentes combinaisons de C et de sigma peuvent conduire à des modèles complexes similaires, avec des performances similaires.
- Lorsque la performance des modèles en fonction des hyperparamètres présente un large plateau est une excellente chose, car cela suggère qu'ils ne sont que peu vulnérables au surajustement dans la sélection du modèle.
- *A contrario*, lorsque les valeurs optimales se présentent sous forme de pics, il est plus difficile de déterminer leur emplacement avec un ensemble de données limité.

3.6/ REVUE SYSTÉMATIQUE DE LA LITTÉRATURE

Nous avons effectué une revue systématique de la littérature courant essentiellement sur les 5 à 10 dernières années. Elle donne un aperçu de l'état de l'intersection du domaine EEG-classification SVM, ainsi que des recommandations tant pour les chercheurs souhaitant passer d'une analyse traditionnelle de l'EEG à des techniques de *machine learning*, que pour les spécialistes du traitement du signal qui voudraient étendre leur domaine de compétence à l'EEG. Ces recommandations méthodologiques devraient favoriser la clarté et la reproductibilité des études.

3.6.1/ MATÉRIEL ET MÉTHODES

“Oui” Des articles de revues scientifiques en anglais, ainsi que des pré-tirages électroniques ont été choisis comme cible de cette revue. La base de données PubMed a été interrogée afin de recueillir une première liste d’articles, à la date du 22 mai 2020.

L’équation de recherché utilisée est la suivante : “eeg support vector machine classification NOT epilepsy NOT seizure NOT review NOT sleep”, ainsi traduite en mots-clefs Mesh : ((“electroencephalography”[MeSH Terms] OR “electroencephalography”[All Fields] OR “eeg”[All Fields]) AND (“support vector machine”[MeSH Terms] OR (“support”[All Fields] AND “vector”[All Fields] AND “machine”[All Fields]) OR “support vector machine”[All Fields]) AND (“classification”[Subheading] OR “classification”[All Fields] OR “classification”[MeSH Terms])) NOT (“epilepsy”[MeSH Terms] OR “epilepsy”[All Fields]) NOT (“seizures”[MeSH Terms] OR “seizures”[All Fields] OR “seizure”[All Fields]) NOT (“review”[Publication Type] OR “review literature as topic”[MeSH Terms] OR “review”[All Fields]) NOT (“sleep”[MeSH Terms] OR “sleep”[All Fields]).

Les documents ont ensuite été inclus ou exclus sur la base des critères énumérés dans le Tableau 3.1.

<u>Critères d’inclusion</u>	<u>Critères d’exclusion</u>
<ul style="list-style-type: none"> • Etudes portant sur la classification des signaux EEG par SVM. • Etudes sur sujets humains uniquement. 	<ul style="list-style-type: none"> • Etudes portant sur l’électrocorticographie (ECoG) et l’EEG intracortical, ou sur la magnétoencéphalographie (MEG). • Revues de la littérature et méta-analyses. • Articles en langues autres que l’anglais. • Etudes portant sur EEG et SVM, mais pas en classification.

Tableau 3.1 – Critères d’inclusion et d’exclusion des articles passés en revue.

L’éligibilité des articles sélectionné a été établie en examinant en premier lieu les titres, puis les résumés, et enfin l’intégralité de l’article. Les documents ne correspondant pas aux critères ont été rejetés.

Un tableau d’extraction de données a été conçu afin de recueillir les éléments pertinents pour nos questions de recherche, sur la base de la revue de la littérature effectuée sur EEG et *deep learning* par Roy *et al.* (2019) [136]. Pour chaque article sélectionné, 36 informations ont été extraites, réparties en catégories : origine de l’article, résumé de l’étude, données utilisées, pré-traitement EEG, méthodologie SVM, résultats obtenus et reproductibilité (Tableau 3.2).

La base de données Pubmed, lorsqu’interrogée, a renvoyé 351 articles, et 4 articles ont été ajoutés de sources supplémentaires. 210 articles ont alors été exclus, et 145 articles ont été inclus pour analyse dans cette revue (Figure 3.7).

<u>Catégorie</u>	<u>Champs</u>	<u>Description</u>
Origine de l'article	Année	Année de publication de l'article
	Journal	Nom du journal dans lequel est paru l'article
	Facteur d'impact	Facteur d'impact du journal de publication
	Pays d'affiliation du 1 ^{er} auteur	Pays dans lequel est situé l'université, le laboratoire ou l'institut de rattachement du 1 ^{er} auteur
Résumé de l'étude	Domaine d'application	Champ technique ou clinique auquel s'applique l'article
	Tâche	Nom ou courte description des tâches effectuées par le sujet lors de l'enregistrement des données.
	Objectif	But principal de l'étude
Données utilisées	Quantité de données	Nombre d'essais par sujet utilisés pour la classification
	Matériel EEG	Modèle de casque EEG utilisé et fabricant
	Nombre d'électrodes	Nombre d'électrodes utilisées pour analyse. Ce peut être différent du nombre d'électrodes présentes sur le casque.
	Fréquence d'échantillonnage	Fréquence d'échantillonnage en hertz utilisée pour l'analyse. Ce peut être différent de la fréquence d'échantillonnage utilisée lors de l'acquisition des données.
	Nombre de sujets	Nombre de sujets utilisés pour analyse. Ce peut être différent du nombre de sujets enregistrés.
Traitement EEG	Pré-traitement	Etapas de traitement appliquées aux données brutes avant l'extraction de caractéristiques
	Gestion des artefacts	Méthode utilisée pour éliminer les artefacts parasitant le signal EEG
	Caractéristiques	Type de caractéristiques extraites et utilisées pour la classification
	Optimisation des caractéristiques	Spécifie si une méthode quelconque a été utilisée pour optimiser l'extraction des caractéristiques
	Nombre de caractéristiques	Taille du vecteur de caractéristiques utilisé pour la classification
Méthodologie SVM	Type de SVM	Méthode de calcul utilisée pour la classification
	Kernel	Fonction noyau utilisée pour la classification
	Optimisation des hyperparamètres	Spécifie si une méthode quelconque a été utilisée pour optimiser les hyperparamètres du kernel
	Hyperparamètres	Hyperparamètres du kernel s'ils sont fournis
	Logiciel	Logiciel utilisé pour la classification
	Gestion des sujets	Analyse intra ou inter-sujets
	Normalisation des données	Méthode de normalisation des données avant classification
Résultats obtenus	Mesure de performance	Métrique utilisée pour juger de la performance de classification
	Procédure de validation	Méthode de division des données pour l'entraînement, le test et l'optimisation du modèle le cas échéant
	Test statistique	Test statistique utilisé pour examiner la performance de classification
	Comparaison des résultats	Comparaison des résultats avec ceux d'autres méthodes utilisées dans l'article ou dans la littérature
	Autres classificateurs	Si comparaison avec d'autres classificateurs dans l'article

Tableau 3.2 – Catégories et description des informations extraites de chacun des articles sélectionnés : origine de l'article, résumé de l'étude, données utilisées, pré-traitement EEG, méthodologie SVM, résultats obtenus et reproductibilité.

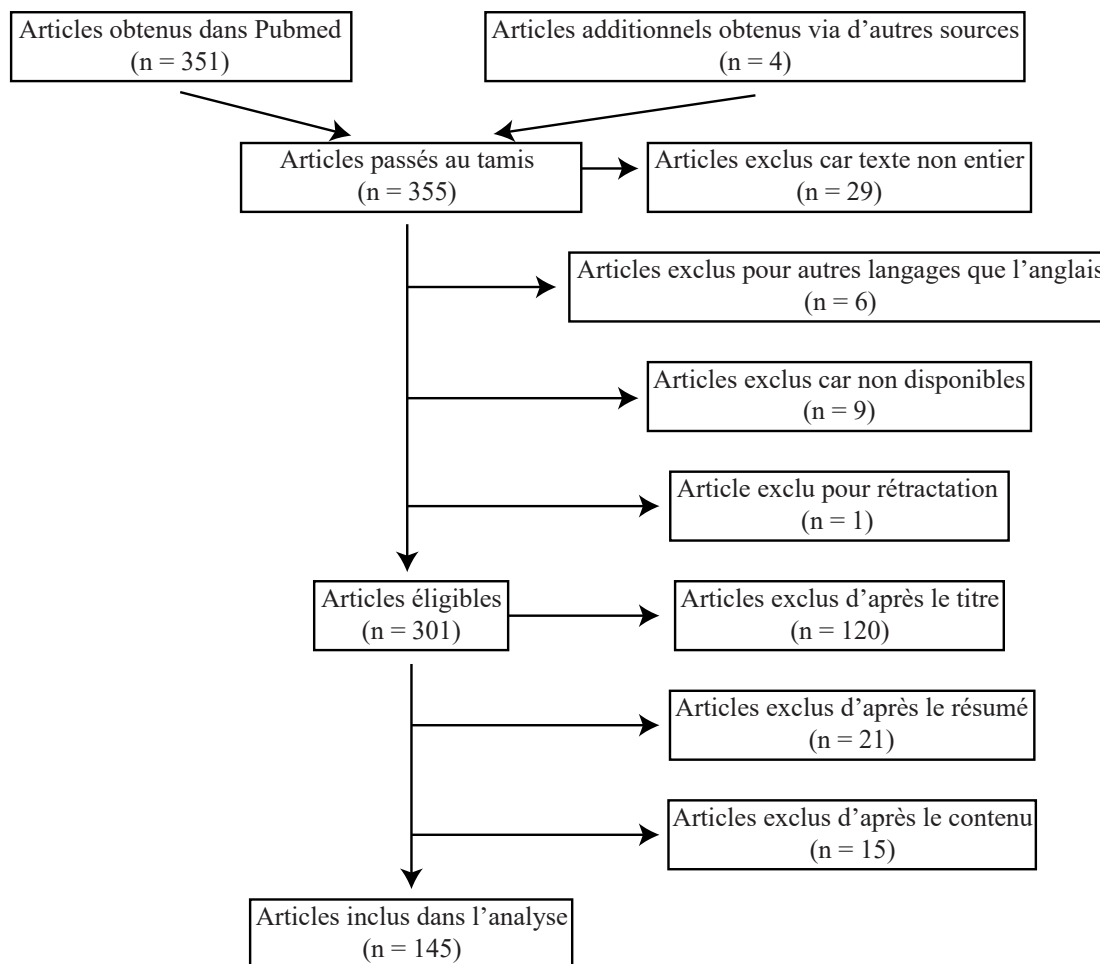


FIGURE 3.7 – Nombre d'articles passés au crible des critères d'inclusion et d'exclusion. Cent quarante-cinq études ont été incluses dans notre analyse.

3.6.2/ RÉSULTATS

3.6.2.1/ ORIGINE DES ARTICLES

Nous avons constaté que la majorité des articles proviennent de Chine et dans une moindre mesure d'Inde et de Turquie (Figure 3.8). De façon surprenante, nous n'avons relevé que 5 articles en provenance des Etats-Unis, alors que nous nous attendions à ce que ce pays figure en tête du classement au vu de sa politique pro-IA.

Les facteurs d'impact des journaux dans lesquels sont publiés ces études sont en moyenne de 3.3 (± 1.9).

3.6.2.2/ RÉSUMÉ DE L'ÉTUDE

La plupart des études se concentrent sur le monitoring de l'activité cérébrale (65 articles), tant émotionnelle [113, 115, 135, 137–158] que cognitive [109, 111, 112, 116, 126, 159–193], tandis que 42 et 34 articles concernent respectivement les domaines de la BCI

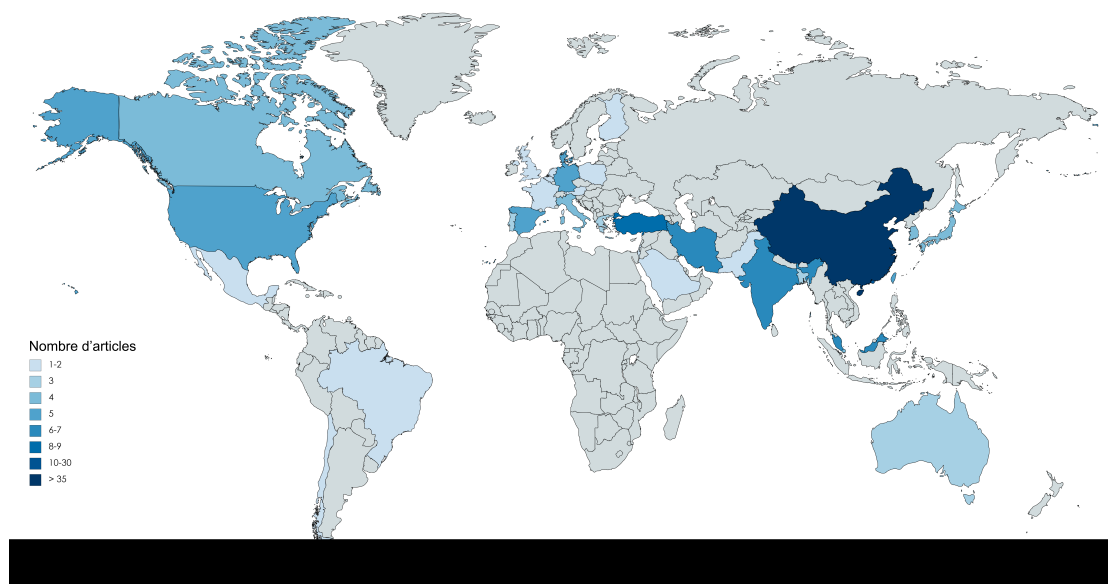


FIGURE 3.8 – Pays d'origine de l'affiliation (laboratoire, institut ou université) du premier auteur. Carte générée à partir de <https://mapchart.net/world.html>.

[117, 133, 134, 194–232] et clinique [110, 112, 114, 233–263] (Figure 3.10). Les quatre autres articles proviennent de champs divers [127, 264–266] (marketing, élimination des artefacts, *etc.*). Cette liste confirme le riche potentiel de l'analyse EEG et motive la nécessité des techniques avancées de traitement des signaux pour aider les cliniciens dans leur interprétation de l'EEG.

Les dates de publication révèlent que ce champ de recherche est récent, avec un démarrage au début des années 2010, et un réel essor à partir de 2015 (Figure 3.9 a), le nombre de publications sur l'année 2018 seule équivalant au nombre d'articles publiés sur 2015 et 2016 combinés.

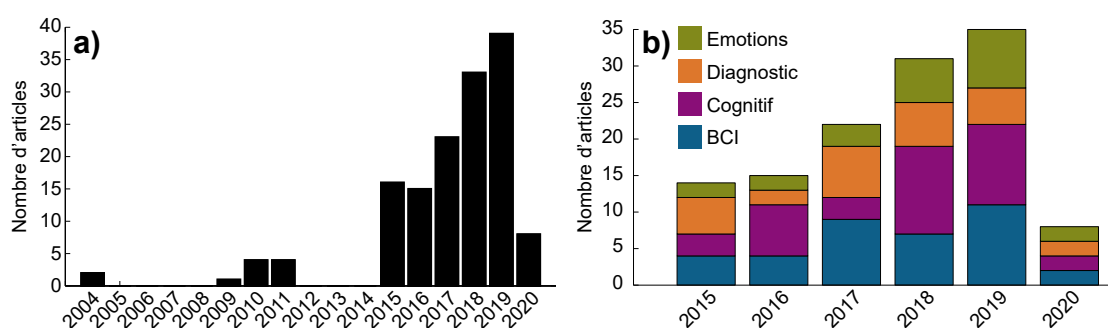


FIGURE 3.9 – Répartition des articles par année et domaine de publication. a) Nombre d'articles publiés par année, de 2004 à mai 2020. b) Nombre d'articles publiés par année et par domaine. Pour faciliter la lecture de l'historgramme, les catégories comprenant moins de 5 publications n'ont pas été incluses dans la figure.

Nous avons également examiné le nombre de publications par domaine par an (Figure 3.9 b). Afin de faciliter la lecture, nous avons concentré l'analyse à partir de 2015 sur les 4 catégories regroupant les plus d'articles, à savoir le diagnostic clinique, les BCI, le monitoring cognitif et émotionnel. Il semble que ces deux dernières catégories soient

celles qui présentent la progression la plus forte.

3.6.2.3/ DONNÉES UTILISÉES

Quantité de données

Le fait qu'une étude porte sur un sujet particulier ou non a un impact sur la performance de classification. Ainsi que nous l'avons vu précédemment, les modèles intra-sujets, qui sont formés et utilisés sur les données d'un seul sujet, conduisent souvent à des performances plus élevées puisque le modèle ne doit tenir compte que de la variabilité intra-individuelle. Cependant, cela signifie que le modèle n'est entraîné sur les données que d'un seul participant, et bénéficie donc le plus souvent d'un nombre d'essais moindre. Dans les études inter-sujets, en particulier les études dites en "*single-trial*", les modèles reçoivent généralement plus de données, car davantage de sujets sont inclus, mais doivent faire face à une plus grande variabilité des données, puisque celle-ci sera non seulement intra- mais aussi inter-individuelle.

Dans les études sélectionnées, 48% se sont concentrées uniquement sur l'approche intra-sujet, 27% sur l'approche inter-sujets, 24% sur l'approche *single-trial*, et 1% (2 articles) ont tenté une combinaison intra/inter-sujets. Il est à noter ici que les études relèvent nécessairement de l'une de ces 3 catégories. Toutefois, la façon dont les données sont gérées sur ce point est rarement mentionnée explicitement dans les études, essentiellement car la méthode employée et la façon de présenter les résultats en dépend, ce qui peut rendre l'information redondante, quoique parfois peu claire. Dans le cadre de cette analyse, nous avons choisi d'examiner les protocoles et les résultats présentés afin d'attribuer l'une de ces catégories aux articles qui ne la spécifiaient pas textuellement (Figure 3.11 a).

Nous nous sommes penchés sur le nombre d'essais utilisés par sujet et par classe (Figure 3.11 b), ainsi que le nombre de sujets (Figure 3.11 c) pour chaque article, selon sa gestion des sujets. Sans surprise, le nombre moyen d'essais est le plus élevé pour la gestion en *single-trial* (694 essais \pm 1637), puisque le nombre de sujets est multiplié par le nombre d'essais enregistrés pour chacun d'eux. Le nombre d'essais consacrés à la classification est le plus bas pour l'intra-sujet (347 essais \pm 581), et est de 415 essais \pm 1089 pour l'inter-sujets. On notera toutefois la grande dispersion du nombre d'essais dans une même catégorie, qui démontre qu'il n'existe pour le moment aucune règle d'or concernant le nombre d'essais minimal à intégrer au classificateur.

De manière intéressante, on remarque par contre une différence plus marquée entre les études ayant un protocole inter-sujets (141 sujets \pm 328) et celles appartenant aux deux autres catégories (13 sujets \pm 10 et 25 sujets \pm 22 respectivement). En effet, le nombre de sujets requis pour classer justement les participants doit être suffisamment élevé pour pouvoir entraîner le SVM, tandis que la classification intra-sujet, surtout lorsqu'elle vise à expérimenter une nouvelle méthode, peut très bien n'être pratiquée que sur un seul sujet.

Paramètres d'enregistrement

Un grand nombre d'articles (voir section Reproductibilité) repose sur des enregistrements EEG privés, effectués expressément pour cette étude. Nous avons donc examiné les paramètres de ces enregistrements, à commencer par le type de casque EEG utilisé (Figure 3.12 a). Nous avons tout d'abord constaté que près d'un tiers des articles ne

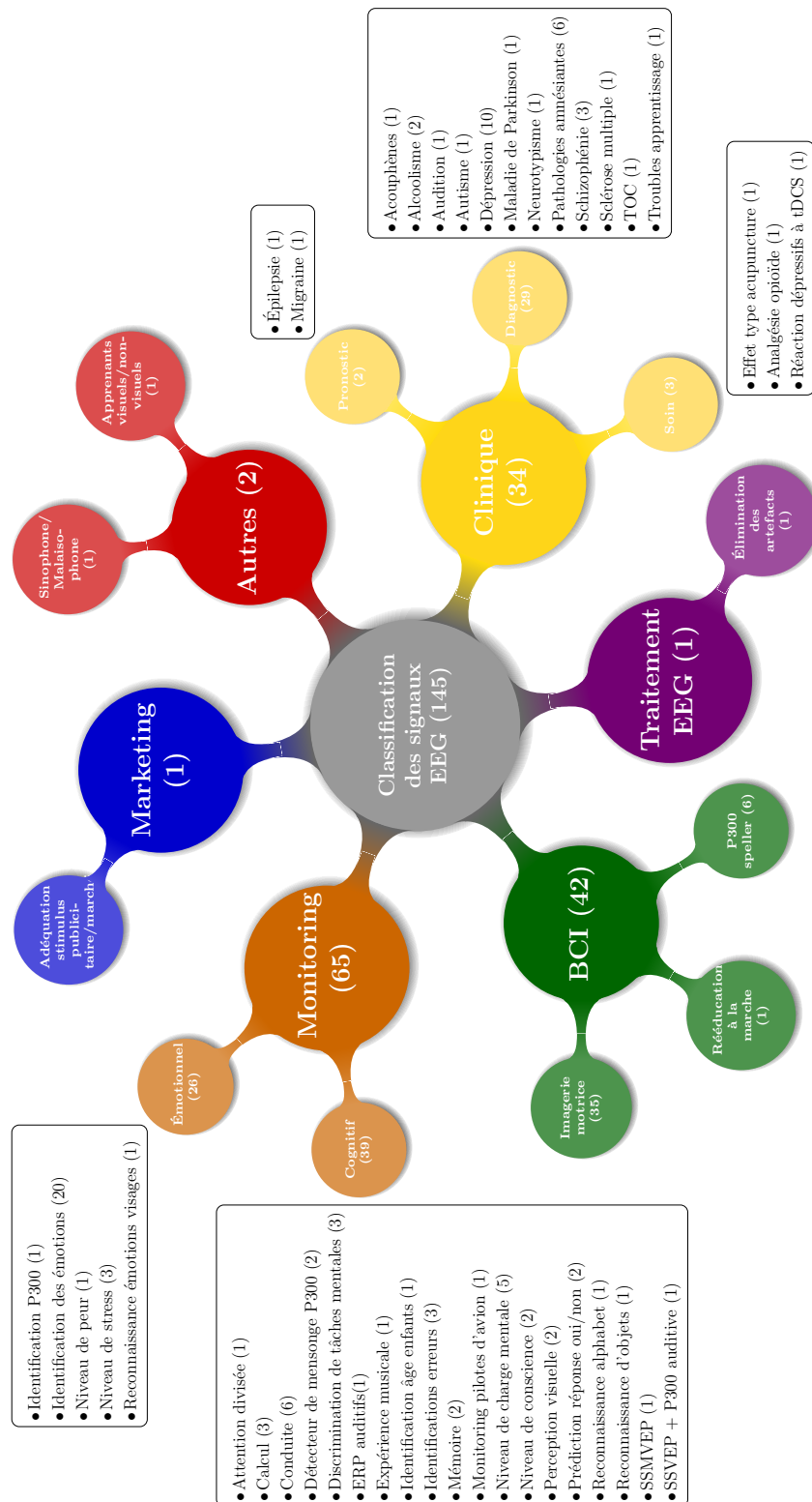


FIGURE 3.10 – Domaine principal de chaque étude. Le nombre d'articles de chaque catégorie est indiqué entre parenthèses. Les études traitant de plus d'un sujet ont été classées selon leur thème principal.

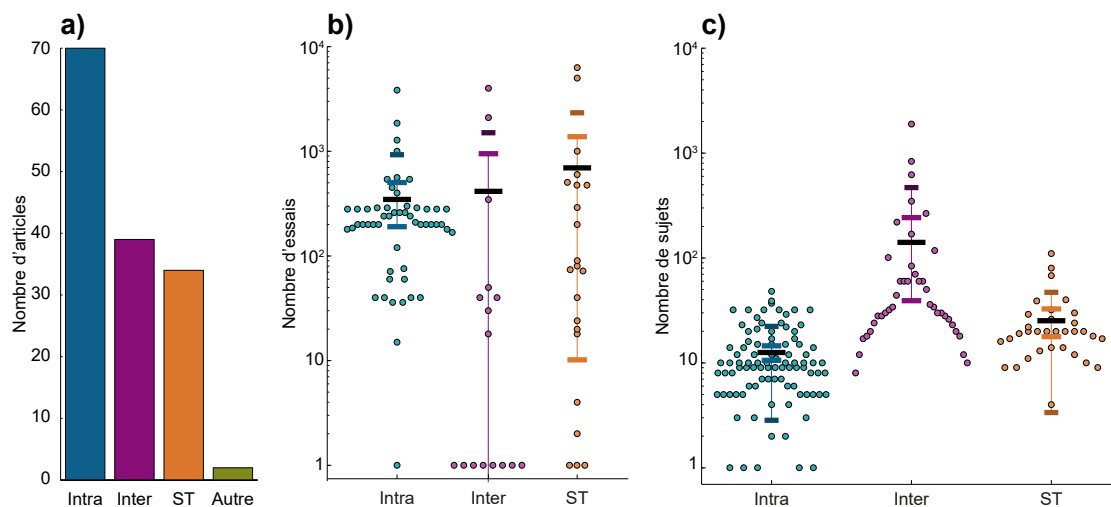


FIGURE 3.11 – Différentes approches sont utilisées dans la gestion des sujets : classification intra-sujet, inter-sujets ou single-trial (ST). a) Nombre de publications par type d'approche. b) Nombre d'essais utilisés par sujet dans chacune des trois approches. c) Nombre de sujets utilisés dans chacune des trois approches.

spécifiaient rien sur le matériel (NS = 47). Un grand nombre de ces articles a eu recours à des jeux de données disponibles publiquement, et ont par conséquent choisi de laisser le lecteur chercher ces informations dans la description desdites données, qu'elle soit fournie ou non avec celles-ci. Lorsque la provenance du matériel est définie, ces appareils proviennent en majorité des entreprises NeuroScan, Biosemi et Electrical Geodesic Inc., qui comprennent jusqu'à 256 canaux. Il est intéressant de constater que les casques EEG portatifs et/ou comprenant peu d'électrodes, tels que les appareils Emotiv, InteraXon et Enobio sont loin de représenter la majorité, alors que la Figure 3.12 b - qui montre le nombre d'électrodes non pas enregistrées mais conservées pour analyse dans les études - marque une nette tendance à l'utilisation de moins de 32 électrodes, autrement dit de l'EEG basse résolution. Les enregistrements EEG sont donc fréquemment faits avec un grand nombre d'électrodes au départ, pour n'en conserver finalement que peu.

De plus, la fréquence d'échantillonnage, qui définit le nombre de points de données pour chaque seconde, est en moyenne de 713 Hz, mais avec un écart-type de 2830 Hz, les valeurs allant de 20 à 20 000 Hz (Figure 3.12 c). Il s'agit ici des fréquences utilisées pour l'analyse et non pour l'enregistrement. Dans quasiment tous les cas, un sous-échantillonnage est pratiqué afin de réduire la masse de données et de faciliter l'analyse en termes de puissance de calcul. Bien que 35.1% des articles aient retenu une fréquence située entre 200 et 256 Hz, on peut constater encore une fois qu'aucune règle d'or n'existe et que le plus souvent, la fréquence de ré-échantillonnage est définie arbitrairement.

3.6.2.4/ TRAITEMENT EEG

Pré-traitement

Le pré-traitement des données EEG comprend généralement plusieurs étapes générales, telles que l'application de filtres passe-bande, le re-référencement des électrodes, l'élimination des artéfacts dus à la pollution électrique de la pièce et aux mouvements

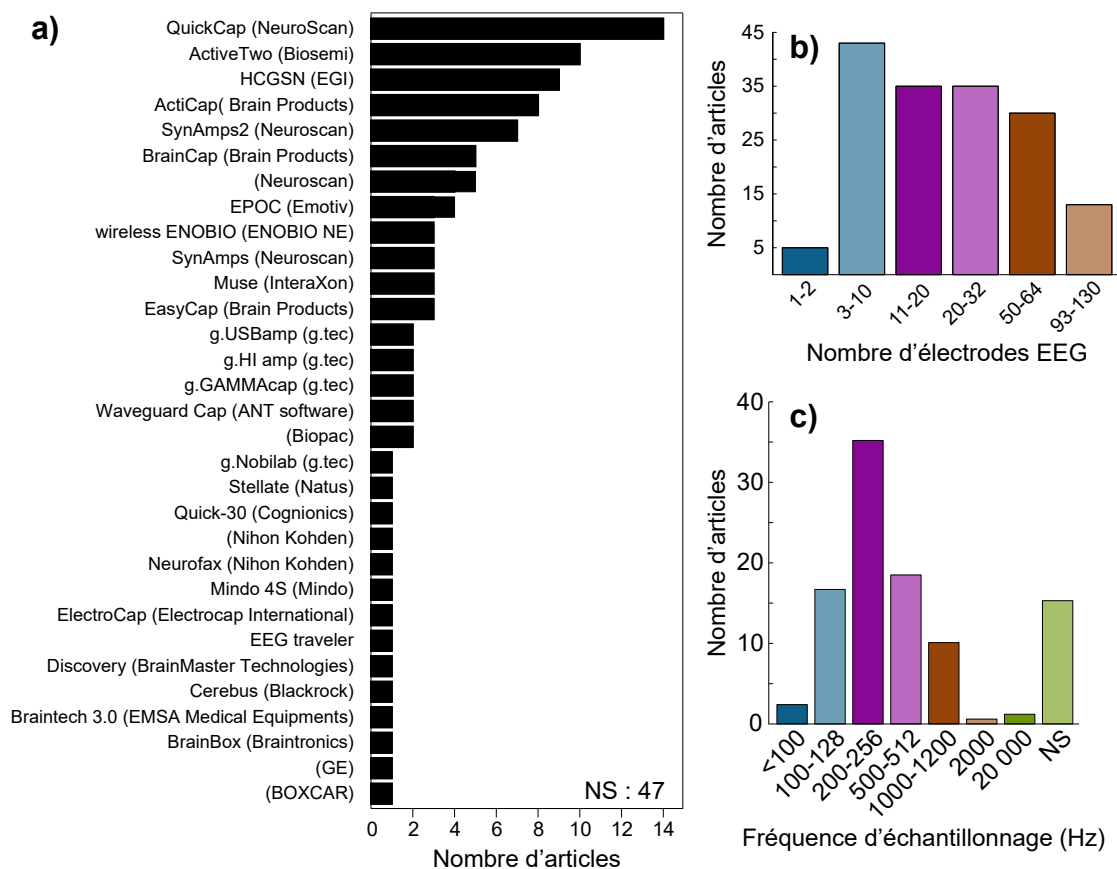


FIGURE 3.12 – Caractéristiques du matériel et fréquence d'échantillonnage. a) Nom et fabricant du matériel EEG utilisé. b) Nombre d'électrodes conservées pour l'analyse. c) Fréquence d'échantillonnage utilisée pour l'analyse (en Hertz).

oculaires et musculaires, la soustraction par rapport à une ligne de base, *etc...* Tout au long du processus d'examen des documents, nous avons constaté qu'un nombre différent d'étapes de prétraitement était utilisé dans les études, dans différentes combinaisons, allant d'un simple filtrage *notch* à l'application de l'ensemble des étapes possibles.

Nous avons séparé les documents en deux catégories selon qu'ils utilisaient ou non des étapes de prétraitement : "Oui", dans les cas où un prétraitement a été utilisé ; "Non", lorsque les auteurs n'ont mentionné aucun prétraitement (Figure 3.13 a).

Élimination des artefacts

Les techniques d'élimination des artefacts sont dédiées au bruit spécifique issu des mouvements oculaires et musculaires du sujet [267].

L'élimination des artefacts est une étape cruciale pour obtenir de bonnes performances de classification EEG [268] comme le souligne l'adage "*garbage in, garbage out*". Nous avons donc examiné les différentes façons de s'en débarrasser (Figure 3.13 b).

Les techniques d'élimination des artefacts nécessitent généralement l'intervention d'un expert humain [269], soit s'appuyant entièrement sur celui-ci, comme dans le cas de l'inspection visuelle (10.3% des articles), soit exigeant son intervention pour la sélection des composants représentant du bruit, comme dans le cas de l'analyse en composantes

indépendantes (ICA) pour 22.8% des études.

Il existe une myriade de techniques de détection automatique des artéfacts, telles que l'utilisation de seuils d'amplitude, la séparation aveugle des sources (BSS), l'élimination du bruit par analyse de variance (xDAWN), *etc...* que nous avons simplement regroupées sous l'appellation "automatique".

Près de la moitié (44.1%) des articles n'ont apparemment pas jugé nécessaire de renseigner une quelconque méthode de nettoyage des artéfacts.

Extraction de caractéristiques

L'extraction des caractéristiques du signal est l'une des étapes les plus exigeantes de la chaîne de traitement EEG traditionnelle. Le principal objectif d'un grand nombre des articles examinés (45.5%) est même d'ailleurs de rechercher les caractéristiques les plus représentatives de l'activité cérébrale des conditions à classer, et des méthodes automatiques de sélection des sous-ensembles de caractéristiques les plus pertinents.

Les informations du domaine fréquentiel de l'EEG sont largement utilisées, notamment les densités spectrales de puissance des bandes de fréquences classiques en EEG, courant de 1 Hz à 40 Hz ont été utilisées comme caractéristiques. Plus précisément, les auteurs ont sélectionné les bandes delta (1-4 Hz), thêta (5-8 Hz), alpha (9-13 Hz), bêta inférieur (14-16 Hz), bêta supérieur (17-30 Hz) et parfois gamma (31-40 Hz). En outre, d'autres études ont utilisé des combinaisons de caractéristiques spectrales, mais aussi temporelles ainsi que l'entropie, entre autres, comme données d'entrée au SVM.

Nous avons ainsi répartis les articles en quatre catégories : "EEG brut", "domaine fréquentiel", "combinaison" (dans le cas où plus d'un type de caractéristiques a été utilisé), et "autre" (pour les articles n'utilisant ni l'EEG brut ni les caractéristiques du domaine fréquentiel). Les études qui n'indiquaient pas le type d'entrée ont été classées dans la catégorie "NS" (non spécifié) (Figure 3.13 c). Nous pouvons remarquer que cette fois, le nombre d'études qui ne spécifient pas ce paramètre est très faible, soulignant donc l'importance cruciale que revêt l'extraction de caractéristiques.

Normalisation

Le SVM suppose que les données avec lesquelles on travaille se situent dans une fourchette standard. Il est donc très important de normaliser les vecteurs de caractéristiques avant de les introduire dans le classificateur pour l'entraînement. Les caractéristiques doivent être normalisées par dimension, et non par essai. Il existe différentes méthodes pour ce faire : faire correspondre les valeurs minimales et maximales d'une dimension donnée à 0 et 1, ou à -1 et 1 (normalisation min-max), ou, le plus souvent, soustraire la moyenne des valeurs et diviser par l'écart type (aussi appelé z-score). Une absence de normalisation peut fausser les résultats du classificateur en donnant davantage d'importance à l'une des caractéristiques, en particulier pour les *kernels* sensibles à cette problématique, tels que le *kernel* RBF.

Or, comme le montre la figure 3.13 d, le nombre d'articles pour lesquels aucune normalisation n'est mentionnée est écrasant : 74.5%. Parmi les articles restants, 13.8% ont opté pour une normalisation z-score, et le reste se répartit sur des variantes de la normalisation min-max.

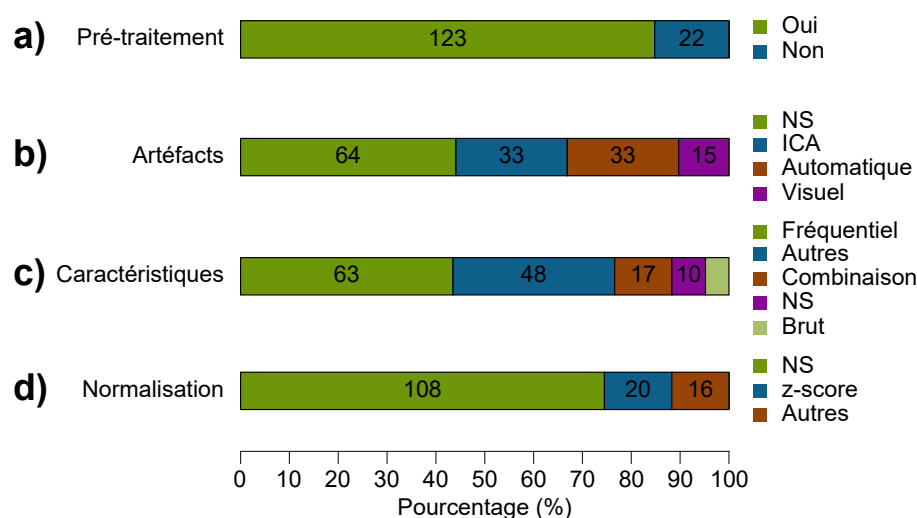


FIGURE 3.13 – Choix effectués concernant le pré-traitement EEG. a) Pourcentage d'articles ayant appliqué un pré-traitement (filtrage, interpolation d'électrodes *etc.*) ou non, ou n'ayant rien spécifié (NS). b) Pourcentage d'articles ayant utilisé une méthode automatique ou humaine, ou n'ayant rien spécifié concernant la conservation, l'élimination ou la correction des artéfacts (NS). c) Pourcentage d'articles ayant utilisé le signal EEG brut, extrait des caractéristiques du domaine fréquentiel, une combinaison des deux, d'autres caractéristiques, ou n'ayant rien spécifié (NS). d) Pourcentage d'articles ayant appliqué une normalisation des données lors de la classification (z-score ou autre), ou n'ayant pas spécifié de normalisation quelconque (NS).

3.6.2.5/ MÉTHODOLOGIE SVM

Ainsi que nous l'avons vu précédemment, le type de *kernel* utilisé et les hyperparamètres choisis pour celui-ci ont une forte influence sur les résultats de la classification SVM. Nous avons donc voulu savoir si l'optimisation de ces paramètres, ou, à défaut, l'indication des valeurs retenues était une pratique courante dans la classification des signaux EEG. Les résultats présentés dans la Figure 3.14 a, b et c montrent clairement que ce n'est pas le cas : sur 145 articles examinés, 103 n'ont usé d'aucune méthode d'optimisation des hyperparamètres, et 112 ne fournissent pas les hyperparamètres retenus. Seules 60 études ont optimisé et/ou fourni les coefficients des *kernels* utilisés. Parmi les protocoles incluant une phase d'optimisation, la méthode retenue dans 66.7% des cas était celle de la recherche par grille.

En outre, 30 articles ne mentionnent pas même le *kernel* utilisé pour entraîner leur modèle SVM (Figure 3.14 d). Nous pouvons tout de même constater que les deux *kernels* les plus populaires sont le linéaire et le RBF. Il nous est en revanche impossible de dire si les auteurs n'ont pas connaissance de l'importance de ces éléments et ne les ont donc pas appliqués du tout, ou s'ils ont utilisé une méthode ou des paramètres particuliers mais n'ont pas jugé bon d'inclure ces informations dans leur article.

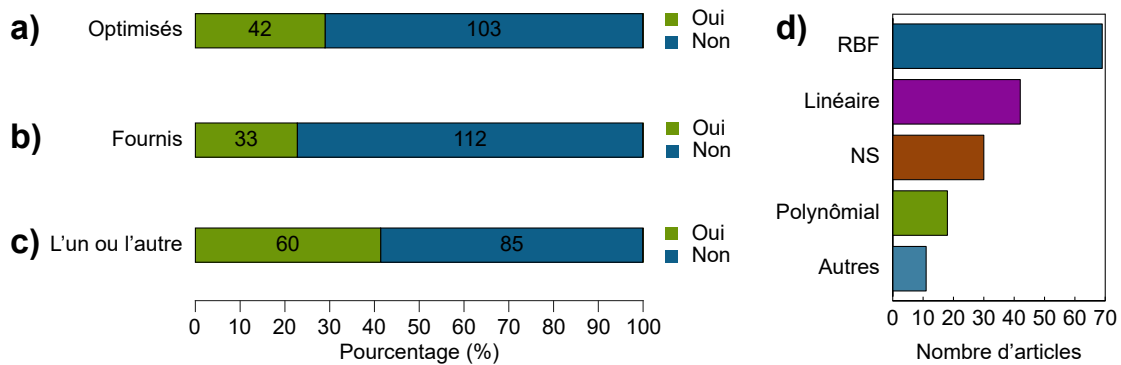


FIGURE 3.14 – Choix effectués concernant les hyperparamètres du SVM. a) Pourcentage d'articles ayant ou non utilisé une méthode d'optimisation des hyperparamètres du SVM. b) Pourcentage d'articles ayant ou non précisé les valeurs retenues pour les hyperparamètres du SVM. c) Pourcentage d'articles ayant ou non proposé l'un ou l'autre. d) Nombre d'articles ayant choisi tel ou tel *kernel* (NS : non spécifié).

3.6.2.6/ RÉSULTATS OBTENUS

Evaluation de la performance de classification

Sans surprise, le moyen de mesure de la performance le plus employé est l'*accuracy* de classification, dans quasiment toutes les études, sauf celles privilégiant l'utilisation de l'aire sous la courbe (AUC). Les autres métriques sont également issues de la matrice de confusion : sensibilité, spécificité *etc...* (Figure 3.15 a).

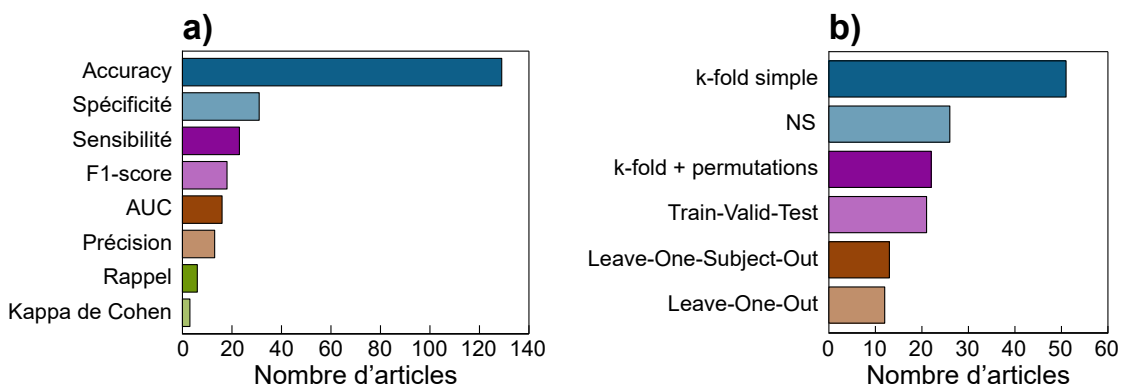


FIGURE 3.15 – Métriques de performance et méthodologie de validation croisée retenues. a) Métriques de performance utilisées dans les articles analysés. Un même article peut avoir employé plusieurs d'entre elles. b) Approche retenue pour la validation croisée (NS : non spécifiée).

Procédure de validation

Les techniques de validation croisée utilisées dans les études sélectionnées sont présentées dans la figure 3.15 b. Là encore, un assez grand nombre d'études (29 articles) n'a pas mentionné l'utilisation de la validation croisée. La catégorie "*k-fold* + permutations" comprend les études qui font des permutations aléatoires entraînement/validation, entraînement/test ou entraînement/validation/test, ainsi que les études qui ont mentionné

la division de leurs données en ensembles de formation, de validation et de test, mais qui n'ont pas fourni de détails sur la méthode de validation. La procédure *k-fold*, avec ou sans permutations, est la plus employée.

3.6.2.7/ REPRODUCTIBILITÉ

Nous avons repris la méthode de Roy *et al.* (2019) [136] pour évaluer la difficulté à reproduire les protocoles mis en œuvre, qui est fonction de la disponibilité des données et du programme utilisés pour la classification :

- Facile : code et données disponibles
- Moyenne : code ou données disponibles sur demande
- Difficile : données disponibles, mais code non disponible
- Impossible : ni le code, ni les données ne sont disponibles. Ainsi que le démontre la figure 3.16, 98.6% des études sont très difficiles ou impossibles à reproduire. C'est un exemple criant de la crise de la reproductibilité.

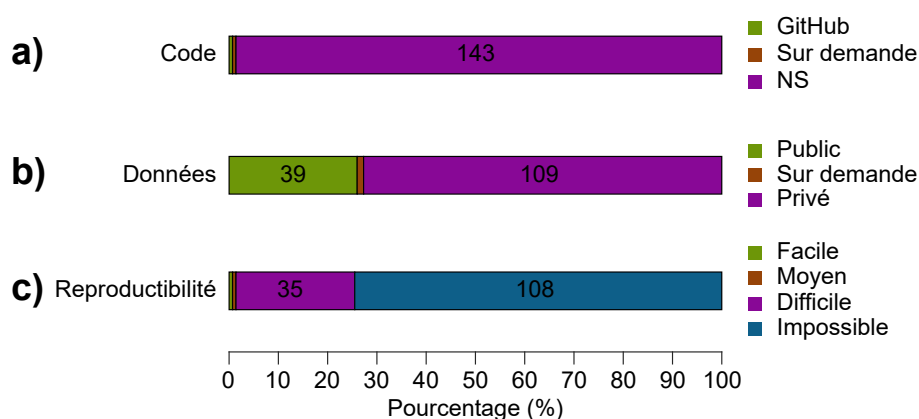


FIGURE 3.16 – Reproductibilité des études examinées. a) Disponibilité du code utilisé pour le pré-traitement et/ou la classification. b) Disponibilité des données EEG utilisées dans les articles. c) Estimation du niveau de reproductibilité des études (Facile : code et données disponibles, Moyen : code ou données disponibles sur demande, Difficile : données disponibles mais pas le code, Impossible : ni les données ni le code ne sont disponibles).

3.6.3/ LIMITATIONS

Nous avons examiné les éléments qui nous semblaient les plus pertinents pour illustrer l'état de l'art de la classification de signaux EEG par SVM. Toutefois, certains paramètres relevés dans les articles de cette revue, tels que la comparaison avec d'autres classificateurs, les logiciels utilisés pour l'analyse, le détail des tâches présentées aux sujets *etc...* n'ont pas été inclus. Ensuite, nous avons choisi de rejeter les articles se focalisant sur les données de sommeil et d'épilepsie, car nous travaillons sur des potentiels évoqués par des tâches précises, et ces données n'en contiennent pas.

Troisièmement, nous n'avons interrogé que la base de données Pubmed, mais il ne fait aucun doute que la consultation d'autres bases, telles que Google Scholar ou Springer

aurait pu nous amener davantage d'articles à prendre en compte. Enfin, l'inclusion d'articles en pré-publication sur la base arXiv aurait pu nous donner un aperçu des méthodes les plus récentes utilisées en 2020, mais nous avons préféré privilégier des articles ayant fait l'objet de relecture par un comité d'experts.

Nous estimons toutefois qu'au vu des résultats assez tranchés que nous avons obtenus, nous présentons ici une synthèse représentative de l'état de l'art.

3.6.4/ RECOMMANDATIONS

Plusieurs recommandations émergent de cette synthèse des travaux existant sur EEG et SVM. Tout d'abord, les données utilisées doivent être clairement décrites : nombre de sujets, d'essais, le nombre de caractéristiques extraites et leur nature exacte, en particulier dans le cas des méthodes de sélection automatiques et de combinaisons de caractéristiques. En outre, lorsque cela est possible, l'utilisation de jeux de données publics est préférable, afin que les méthodes d'extraction et de classification puissent être aisément comparées, ou à défaut, mettre à disposition les enregistrements "maison".

Concernant la comparaison des méthodes, le partage du programme utilisé, incluant les choix effectués en termes d'hyperparamètres, de *kernels* et de détails du modèle, est absolument nécessaire pour la vérification par les pairs et pour la reproductibilité des résultats. L'architecture de la validation croisée doit également être clairement spécifiée, puisque des erreurs à ce niveau peuvent mener à un surajustement du modèle, et à une performance artificiellement meilleure.

Enfin, une comparaison systématique des résultats obtenus par de nouvelles méthodes doit être réalisée avec les méthodes bien connues dans l'état de l'art, ou à défaut, avec des données générées aléatoirement afin de s'assurer que les performances obtenues ne sont pas dues au hasard.

3.6.5/ CONCLUSION

Nous avons donc, tout au long de ce travail de thèse, pris en compte autant que possible ces recommandations et présenté des travaux à la fois clairs et reproductibles sur la classification de signaux EEG dans le cadre du diagnostic différentiel entre la maladie d'Alzheimer et la dépression d'une part, et dans le diagnostic des troubles de la conscience liés à un traumatisme hypoxique d'autre part.



DÉPRESSION ET MALADIE D'ALZHEIMER : LE PROJET RIHANNA

LA MÉMOIRE MUSICALE COMME CIBLE DU DIAGNOSTIC DIFFÉRENTIEL ENTRE DÉPRESSION ET DÉMENCE

MUSIQUE ET MÉMOIRE

Les difficultés mnésiques de la maladie d'Alzheimer et de la dépression relevant de mécanismes *a priori* différents [270], la mémoire semble être un axe de recherche pertinent pour des neuromarqueurs discriminants.

Parmi les fonctions mnésiques, la mémoire sémantique musicale, définie par la capacité d'un sujet à identifier des extraits musicaux connus, sera privilégiée. Il s'agit d'une sous-catégorie de la mémoire sémantique, soit l'encodage des concepts indépendamment des circonstances de l'apprentissage. Elle permet l'identification des musiques entendues et induit un sentiment de familiarité lorsque celles-ci sont bien connues du sujet, autrement dit la reconnaissance d'une musique sans pour autant y associer systématiquement une étiquette verbale. L'association de cette représentation perceptive à des connaissances verbales, comme par exemple le titre ou l'interprète de la musique, implique cette fois l'intercession de la mémoire épisodique, qui permet la mémorisation des informations autobiographiques contextuelles aux conditions spatiales et temporelles d'exposition à la musique [271].

Bien que la musique fasse intégralement partie de la culture humaine depuis les débuts de l'espèce, la mémoire musicale n'a été que peu étudiée par rapport aux formes de mémoire visuelle et verbale dans le vieillissement et la démence. Au cours de ces dernières années, toutefois, quelques études cliniques, portant sur des patients auparavant musiciens atteints de la maladie d'Alzheimer, ont mis en évidence une préservation de la capacité musicale, contrairement aux déficits mnésiques et langagiers [272], et ont conduit à des travaux plus systématiques chez les patients Alzheimer non musiciens.

Étant donné que la pathologie Alzheimer affecte les lobes temporaux dès le stade précoce, il serait attendu que la mémoire musicale à long terme explicite, tant sémantique qu'épisodique soit altérée très tôt, tandis que les régions cérébrales impliquées dans la mémoire implicite semblent épargnées jusque dans les stades les plus avancés, suggérant une conservation des capacités procédurales des musiciens [273]. Toutefois, des études de cas [274–276] ont démontré une certaine préservation de la mémoire sémantique dans la reconnaissance de mélodies familières ainsi que des traces d'apprentissage de nouvelles chansons aux différents stades de progression de la dégénérescence.

Ainsi, deux tâches d'apprentissage implicite en mémoire sémantique musicale ont été menées par Moussard *et al.* (2008) [277] chez les sujets âgés sains et ceux atteints

de maladie d'Alzheimer à un stade léger et modéré : une tâche de catégorisation (appréciation ou non d'extraits joués au piano) et une tâche de segmentation de flux (sons à définir comme déjà entendus ou non), dans lesquelles les patients Alzheimer présentaient des performances équivalentes aux groupes contrôles. La mesure de la mémoire musicale épisodique, après apprentissage explicite par discrimination de nouvelles mélodies, a montré que la performance des patients Alzheimer était inférieure à celle des contrôles, indiquant une atteinte précoce de la mémoire épisodique dès le stade léger de la maladie. Ces résultats rejoignent ceux de l'étude de Bartlett *et al.* (1995) [278] sur la reconnaissance musicale chez des patients Alzheimer à un stade modéré, au cours de laquelle les patients Alzheimer ont commis le nombre plus important de fausses alarmes pour les mélodies familières et non familières, confirmant ainsi un déficit dans le processus d'encodage épisodique.

D'un point de vue neurofonctionnel, l'implication d'un réseau neural plus vaste de la mémoire sémantique musicale a été démontrée, avec des activations des aires cingulaires antérieure et moyenne, du *gyrus* frontal inférieur gauche (aire de Broca), et plus généralement de régions temporales et préfrontales dans les deux hémisphères, impliquées spécifiquement dans la mémoire sémantique musicale mais pas verbale, ce qui pourrait contribuer à la résistance parfois extraordinaire de cette mémoire dans la maladie d'Alzheimer [279]. D'autre part, par comparaison en IRMf des régions cérébrales impliquées dans le traitement des musiques bien connues avec l'anatomie de la progression des altérations de la maladie d'Alzheimer, [280] ont montré que les réseaux concernés n'étaient pas les mêmes : les régions du *gyrus* cingulaire antérieur et de la pré-aire sensorimotrice impliquées dans la mémoire musicale à long terme sont les dernières touchées par la maladie d'Alzheimer et ne recourent quasiment pas celles touchées par la dégénérescence du cortex. Cette explication neuro-anatomique soutient l'hypothèse de la préservation de la mémoire musicale chez les patients atteints de la maladie d'Alzheimer, et ce même à un stade avancé.

La mémoire à long terme pourrait donc être principalement évaluée en proposant des études en mémoire sémantique musicale, dans le but de comprendre l'organisation des connaissances musicales en mémoire.

L'ensemble de ces études montre une préservation de la capacité de familiarisation aux nouvelles mélodies chez les patients Alzheimer. De plus, ce sentiment est plus fort pour les mélodies suscitant des émotions intenses (ex. joie et peur), avec la mise en évidence d'une mémorisation plus importante des extraits musicaux enjoués associés à des émotions positives. Autrement dit, les émotions suscitées par la musique augmenteraient l'encodage et la consolidation des traces mnésiques [281]. À l'heure actuelle, cependant, aucune étude portant sur la mesure de la mémoire musicale n'a été effectuée chez les patients dépressifs.

Néanmoins, il a été noté que ces patients interprètent plus facilement les mélodies dans un sens défavorable, quel que soit le nombre d'expositions préalables ceci étant lié à un biais de négativité [282]. Cet effet de congruence à l'humeur nous permet de supposer que la modification des capacités d'apprentissages implicites portera essentiellement sur les mélodies aux contenus émotionnels positifs tandis que ces capacités resteront relativement bien préservées pour les mélodies aux contenus émotionnels négatifs.

IMAGERIE AUDITIVE MUSICALE

La plupart des études d'évaluation de la mémoire musicale nécessite une participation active des sujets. Cependant, les patients atteints de démence souffrent fréquemment de troubles de compréhension et d'expression qui peuvent alors biaiser les résultats obtenus. Or, il est possible grâce à l'imagerie cérébrale de mesurer la mémoire musicale en l'absence de toute instruction, grâce tout d'abord à un apprentissage implicite par simple exposition à des musiques non familières, puis à la mesure de la rétention mnésique par amorçage implicite. L'apprentissage implicite est l'acquisition de connaissances par le système cognitif, sans intention ni conscience manifestées par les sujets [283].

Plusieurs études ont pu mesurer l'activation de la mémoire musicale de manière automatique, c'est-à-dire avec une participation minimale du sujet, ce qui permet d'éviter tout problème concernant des capacités de réponse altérées. Pour cela, les auteurs sont partis de la capacité du cerveau à provoquer la sensation d'entendre des sons, et ce même en absence de stimulation auditive. Cette illusion a été décrite en tant qu' "imagerie auditive" dans la littérature. Il a été démontré par Kraemer *et al.* (2005) [284] en IRM fonctionnelle que l'écoute passive d'extraits musicaux dans lesquels ont été introduites des périodes de silence à la place du son était suffisante pour provoquer une image auditive lors de ce silence, ainsi que le caractère obligatoire de ce phénomène. Le cerveau a ainsi tendance à compléter de manière automatique les blancs lors de l'écoute de chansons familières. De plus, une différence d'activité neuronale durant ces silences a été observée entre les musiques définies comme inconnues et celles évaluées comme familières par les auditeurs, et ce, que les pièces soient purement instrumentales ou qu'elles comportent des paroles. En outre, une activation neuronale des cortex auditifs primaire et secondaire lors de l'écoute de chansons inconnues ou familières a été mise en évidence.

L'imagerie musicale implique des représentations mentales d'une complexité similaire à celle de la musique perçue, et est aussi limitée par les mêmes restrictions : la musique possède un ordonnancement temporel, et la partie suivante ne peut être entendue ou imaginée que s'il y a eu activation de la précédente. Seule l'imagerie musicale est unidirectionnelle [285]. La capacité à imaginer une musique progressant dans le temps et de scanner cette progression lors de l'écoute suggère le maintien d'un métronome interne qui donne le tempo de la musique interne. Ce processus peut être contrôlé puisque l'on peut imaginer la même musique à un tempo plus ou moins rapide. Toutefois, le tempo auquel la musique aura le plus souvent été entendue est stocké dans la mémoire à long terme et est très fiable. Ce métronome interne permet à l'imagination d'activer des représentations mentales qui évoluent correctement dans le temps. Ainsi, lors de l'insertion d'un silence dans la musique, la complétion par le cerveau se fera sur le même tempo, permettant, lors de la reprise de la musique, d'en être exactement au même point dans la perception et dans l'imagerie invoquée [286].

Dans cette étude, le chronomètre interne est contrôlé par la donne d'un stimulus auditif perceptible qui fournit sa structure temporelle comme une instruction au sujet pour l'imagerie musicale. L'imagerie musicale qui se produit pendant les périodes de silence dans la musique familière apparaît donc comme la continuation spontanée d'une représentation de la mémoire à long terme. Il s'agit donc d'une tâche implicite, avec comme hameçon les passages musicaux entre 2 silences, contrairement à une tâche où il est demandé au sujet d'imaginer activement une musique, correspondant à une imagerie explicite.

La perception mentale d'une musique familière apparaît donc particulièrement exacte, tant au point de vue des paroles que du tempo [285, 287], de la hauteur [288, 289], du timbre [290] et des émotions véhiculées [291].

Afin d'éviter les problèmes de coût et de disponibilité de l'appareillage, le protocole IRMf de Kraemer *et al.* (2005) [284] a ainsi été transposé vers l'électroencéphalographie haute résolution, non invasive et transportable au lit du patient. L'analyse de ce phénomène en EEG-HR repose sur les potentiels évoqués auditifs observables sur le tracé de l'électroencéphalogramme après une stimulation auditive répétitive. L'excellente résolution temporelle caractéristique de l'EEG rend en outre cette méthode particulièrement adaptée pour l'étude des stimuli évoluant rapidement dans le temps tels que la musique.

En 2016, cette méthode d'investigation, nommée "projet Rihanna", a été validée en phase préclinique par Gabriel *et al.* (2016) [292] sur vingt sujets sains de moins de 60 ans. L'étude consistait à faire écouter durant un premier EEG 5 chansons bien connues du sujet et 5 qu'il ne connaissait pas du tout, puis dans un second temps à faire apprendre par une exposition régulière (en moyenne deux fois par jour pendant quinze jours) ces chansons inconnues (Figure 3.17). Au terme de cette phase d'apprentissage, un second EEG était pratiqué afin de mesurer la mémorisation. Vingt périodes de silence de deux secondes chacune étaient insérées dans les chansons lors de l'examen EEG, et les réponses neuronales des sujets ont été examinées durant ces arrêts du stimulus musical (Figure 3.18).

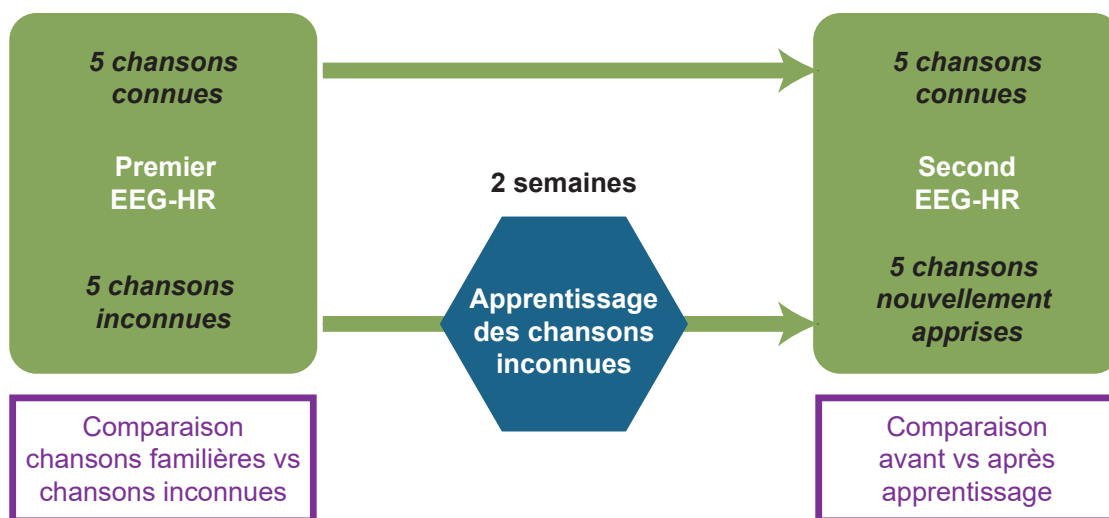


FIGURE 3.17 – Déroulement de la procédure expérimentale.

RÉSULTATS EN ERPS

L'analyse de la mémoire musicale par imagerie auditive en EEG-HR repose sur les potentiels ERPs auditifs observables sur le tracé électroencéphalographique après une stimulation auditive répétitive. Une segmentation de l'activité à court terme sur les volontaires sains a mis en évidence l'apparition successive de cinq périodes d'intérêt à partir du début de la période de silence.

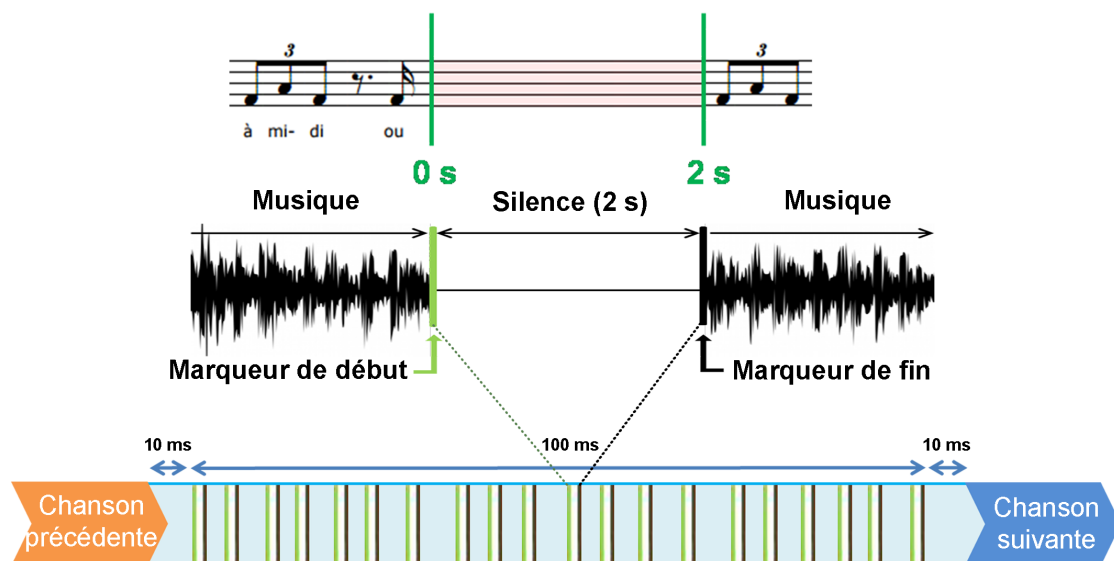


FIGURE 3.18 – Schéma représentant le déroulement de l'expérience EEG. Le début et la fin de chaque silence est repéré par un marqueur différent afin de retrouver ces périodes sur le tracé EEG. 20 silences d'une durée de 2 secondes ont été introduits, avec un délai variant de 2.5 à 3.5 s entre les périodes de silence. Les 10 chansons sont présentées à l'écoute de manière aléatoire, deux fois de suite. Des pauses sont proposées au sujet entre les sessions.

Réponse cérébrale précoce : N100-off

Dans l'étude de Gabriel *et al.* (2016) [292], aucune différence significative n'a été notée entre les 4 catégories considérées au niveau de la N100-off, ce qui signifie qu'elle n'est pas liée à la familiarité de chansons. Pour cela, on pourrait supposer que cette carte correspond à un traitement attentionnel automatique du cerveau lors de l'arrêt brusque de la chanson. Ceci est confirmé par l'origine frontale de cette onde, qui est notamment impliquée dans les processus attentionnels [293].

Mécanismes attentionnels : MMN et P3a

- La MMN est ici induite par un silence inattendu [294]. Son rôle est de signaler une rupture entre le stimulus sensoriel et les représentations mentales stockées dans la mémoire [295]. Son amplitude varie en fonction de la puissance de la violation mnésique. L'amplitude de la MMN est significativement plus petite pour les chansons inconnues que pour les autres chansons. Cette diminution est reflétée par une diminution de l'activité dans le cortex préfrontal et l'aire motrice supplémentaire (SMA) [284]. Le cortex préfrontal est une région liée aux processus de détournement automatique de l'attention et son activité était réduite de façon similaire pour les chansons inconnues et nouvellement apprises. Cette réduction n'est pas perceptible à l'enregistrement de surface car la plus grande partie du signal provient de la SMA, où le signal n'est réduit que pour les chansons inconnues. Cette structure est considérée comme un générateur de la MMN [296] et est fortement activée pendant la perception [297] et l'illusion [298] de chansons familières.
- Évoquée en réponse à des stimuli rares ou nouveaux, la P3a, à distribution fronto-

centrale, semble correspondre au troisième train d'ondes ERP observé. D'après [299], la P3a survient en général lors d'une violation des attentes musicales, c'est-à-dire lorsque des accords incongrus apparaissent dans la mélodie. Dans cette étude, c'est l'absence de musique au beau milieu d'une séquence musicale qui génère l'incongruité provoquant la P3a. Cette réorientation de l'attention semble être ici indépendante de la familiarité, en lien avec une égale activation de la SMA, responsable de l'effet de familiarité, dans toutes les conditions.

Réponses cognitives : P300 et N400

- La première partie de cette période de stabilité est reflétée par une activité pariétale positive. Il s'agit d'une onde P300, connue pour son implication dans les processus mnésiques et l'apprentissage [300, 301]. Une P300 de plus grande amplitude est observée pour les chansons connues, confirmant les résultats de [302] dans lesquels l'amplitude de la P300 augmente lorsque les stimuli sont mieux encodés dans la mémoire.
- La carte fronto-centrale négative observée chez Gabriel *et al.* (2016) [292] semble correspondre à l'effet de familiarité de la N400. Toutefois, bien qu'il ait été montré que l'amplitude de la N400 diminue de façon inversement proportionnelle à la familiarité [303, 304], il n'a pas été observé ici de corrélation entre le degré de familiarité et la force de la réponse.

RÉSULTATS DANS LE DOMAINE FRÉQUENTIEL

Une analyse de la puissance spectral a été effectuée sur les neuf ROI afin d'évaluer les mécanismes cérébraux impliqués dans les activités à long terme pendant la période de silence. Une différence significative a été mise en évidence entre les chansons familières et inconnues dans la ROI centro-médiane, principalement entre 5 et 10 Hz, correspondant à la fois à la bande thêta supérieure (5-7 Hz) et à la bande alpha inférieure (8-10 Hz).

Les chapitres 4 et 5 se basent sur ces résultats, publiés par Gabriel *et al.* (2016) [292], dont l'article complet est disponible dans l'annexe A.

MODULATION DE L'ENCODAGE ET DU RAPPEL EN MÉMOIRE PAR LES STIMULI AFFECTIFS

Afin d'étudier la mémoire musicale dans des conditions optimales, il est important de tenir compte du fait que la mémoire n'est pas qu'une somme de souvenirs, mais obéit à des affects personnels, tels que les émotions ou le caractère plaisant ou non des stimuli à mémoriser. Ainsi l'apprentissage, et par conséquent ces neuromarqueurs, peut être influencé par ces paramètres, dont il est alors nécessaire de tenir compte. Une des difficultés de cette étude est de comprendre comment les auditeurs réagissent à la musique. Il a souvent été suggéré que l'attraction exercée par la musique vient de son pouvoir émotionnel. De plus, du fait de sa grande ubiquité, la musique a été identifiée comme un élément important dans la construction des souvenirs autobiographiques.

4.1/ MUSIQUE, APPRENTISSAGE ET APPRÉCIATION

La musique évoque généralement des émotions grâce à la manipulation des attentes cognitives, et la musique plaisante l'est souvent parce qu'elle construit et répond à ces attentes [305]. Les préférences musicales sont connues pour être associées aux réponses émotionnelles [306]. Il sera donc nécessaire d'explorer l'influence de l'appréciation musicale sur l'apprentissage.

Une des plus importantes motivations à écouter de la musique est le plaisir induit par cette activité. Bien que de nombreux déterminants tels que la personnalité de l'individu [307], ou l'information contextuelle contribuent à la préférence musicale, la structure de la chanson joue un rôle prédominant. Si un individu apprécie d'entendre une séquence de sons organisée dans le temps, ce n'est pas uniquement à cause des propriétés acoustiques de ces sons (timbre, harmoniques), mais aussi parce que leur organisation en schémas génère des attentes et des anticipations de la part du cerveau [308].

Les attentes et les anticipations sont nécessaires pour suivre l'information musicale entrante à de multiples niveaux [309] et pour faire des prédictions sur ce qui va être entendu juste après [310–312]. Elles sont générées en temps réel, sous la forme d'un composant semblable à l'imagerie inclus dans l'écoute de la musique. À cet égard, la dimension temporelle est la clef de la compréhension des mécanismes neuraux menant à la préférence musicale [309, 310].

Comme nous sommes capables de recréer de façon interne une expérience auditive du stimulus musical, les processus spatiaux et temporels mis en place durant l'imagination d'une chanson sont remarquablement similaires à ceux observés pendant la perception [298, 313–315]. La structure musicale (tempo, hauteur), mais aussi les émotions intégrées dans une chanson sont mentalement perçues avec une grande efficacité [291]. Ainsi, la préférence musicale devrait être un autre paramètre préservé pendant l'imagerie musicale, ce qui rend nécessaire sa prise en compte lors de l'étude des ERPs, puisque tout comme les émotions, elle est susceptible d'en influencer l'amplitude ou la latence.

4.2/ RÉSUMÉ DE L'ARTICLE

Les résultats de cette étude ont été publiés dans la revue *International Journal of Psychophysiology* en 2018, sous le titre *Evidence for a neural signature of musical preference during silence*. L'article en lui-même figure dans l'annexe B.

4.2.1/ INTRODUCTION

Bien que l'une des plus grandes motivations pour écouter de la musique soit le plaisir suscité par cette activité, on sait peu de choses sur les mécanismes neuronaux qui suscitent notre goût pour une chanson. La capacité à faire des prédictions sur la continuation d'une chanson tout en l'écoutant est un déterminant important de l'appréciation musicale, mais elle est difficile à explorer car le traitement neural de la musique en cours se produit en même temps que les processus de prédiction sont générés. Pour éviter cette limitation, il est possible d'explorer les mécanismes de prédiction sous-jacents du processus de préférence musicale, non pas pendant la perception, mais pendant l'imagerie musicale spontanée due à la complétion automatique de paroles manquantes.

4.2.2/ MATÉRIEL ET MÉTHODES

Nous avons exploré les mécanismes neuronaux de l'appréciation musicale chez vingt volontaires sains. Dans un premier temps, nous leur avons demandé d'écouter cinq chansons inconnues pendant deux semaines. Par la suite, les deux chansons jugées les plus appréciées et les deux moins aimées ont été sélectionnées.

Les corrélats neuronaux de la préférence musicale ont été étudiés par EEG-HR à l'aide d'un casque Geodesic Sensor Net à 256 électrodes pendant l'écoute de ces chansons dans lesquelles vingt périodes de silence de deux secondes chacune avaient été insérées.

4.2.3/ RÉSULTATS ET DISCUSSION

Plusieurs mécanismes neuronaux liés à l'évaluation ont été mis en évidence. Tout au long de la période de silence, des activités oscillatoires spécifiques sont apparues à l'analyse temps-fréquence, entraînant des changements dans l'activité thêta et alpha au sein de structures cérébrales particulières telles que le gyrus frontal inférieur et le gyrus temporal supérieur, connus pour être impliqués dans la récupération de la musique. Cette étude

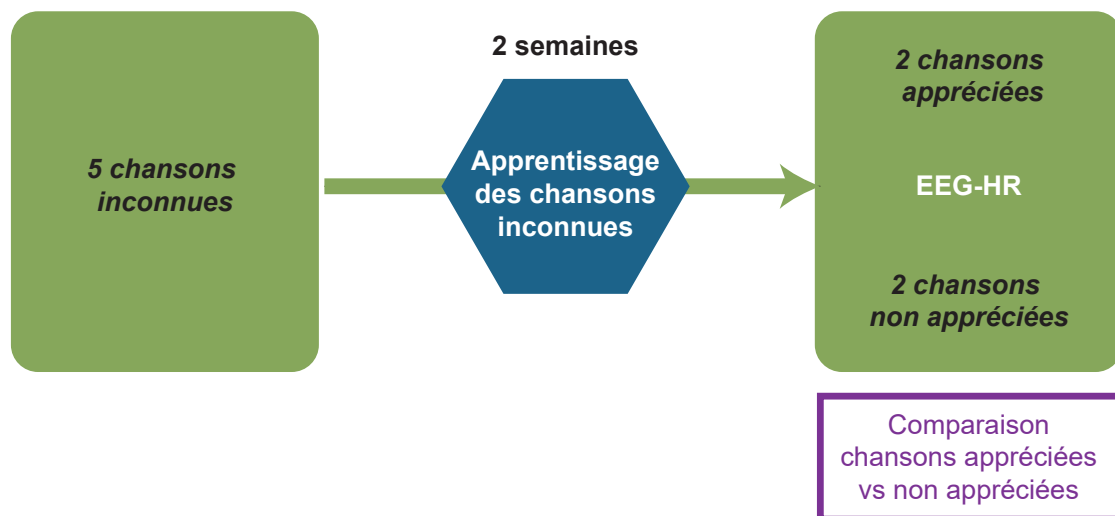


FIGURE 4.1 – Déroulement de la procédure expérimentale.

montre comment notre connaissance du stimulus constitue la base d'un jugement de préférence.

4.3/ EVIDENCE FOR A NEURAL SIGNATURE OF MUSICAL PREFERENCE DURING SILENCE

Joucla Coralie^{1,2}, Nicolier Magali^{1,2,3}, Giustiniani Julie^{1,3}, Brunotte Gaele¹, Noiret Nicolas^{4,5}, Monnin Julie^{1,2,3}, Magnin Eloi^{4,6}, Pazart Lionel^{1,2}, Moulin Thierry^{1,2,6}, Haffen Emmanuel^{1,2,3}, Vandel Pierre^{1,2,3,4}, Gabriel Damien^{1,2}

¹ Centre d'investigation Clinique-Innovation Technologique CIC-IT 1431, Inserm, CHRU Besançon, F-25000 Besançon, France

² Neurosciences intégratives et cliniques EA 481, Univ. Franche-Comté, Univ. Bourgogne Franche-Comté, F-25000 Besançon, France

³ Service de psychiatrie de l'adulte, CHRU Besançon, F-25000 Besançon, France

⁴ Centre Mémoire de Ressource et de Recherche de Franche-Comté, CHRU Besançon, F-25000 Besançon, France.

⁵ Laboratoire de psychologie EA 3188, Université de Franche-Comté, F-25000 Besançon, France

⁶ Service de neurologie, CHRU Besançon, F-25000 Besançon, France

4.3.1/ ABSTRACT

One of the most basic and person-specific affective responses to music is liking. The present investigation sought to determine whether liking was preserved during silence in music. To this purpose, we inserted two-second silent intervals into liked and disliked songs, a method known to automatically recreate a mental image of these songs. Neural correlates of musical preference were measured by high-density electroencephalography

in twenty subjects who had to listen to a set of five pre-selected unknown songs the same number of times for two weeks. Time frequency analysis of the two most liked and the two most disliked songs confirmed the presence of neural responses related to liking. At the beginning of silent intervals (400-900 ms and 1000-1300 ms), significant differences in theta activity were originating from the inferior frontal and superior temporal gyrus. These two brain structures are known to work together to process various aspects of music and are also activated when measuring liking while listening to music. At the end of silent intervals (1400 – 1900 ms), significant alpha activity differences originating from the insula were observed, whose exact role remains to be explored. Although exposure was controlled for liked and disliked songs, liked songs were rated as more familiar, underlying the strong relationship that exists between liking, exposure, and familiarity.

Keywords: musical preference, auditory imagery, high density EEG

4.3.2/ INTRODUCTION

Sometimes it's the space between the notes that makes all the difference in music. Silence in music conveys information of major importance that was efficiently used by composers like Mozart who said that "music is not in the notes, but in the silence between". Some authors have suggested that high-level musical information, such as familiarity, is also present during silence: "When you listen to Mozart, the silence that follows is still Mozart" [316].

In the early 2000s, neuroimaging studies confirmed that familiarity was processed during silences, even when these silences were artificially inserted into music [284, 292]. Two seconds of silent periods inserted in passively listened familiar music generate a neural activity within regions of secondary auditory cortex, of the dorsolateral prefrontal cortex and the supplementary motor area can occur [284]. In electroencephalography (EEG), the response to familiar music is reflected by an increase in theta power during the whole silence period [292]. Interestingly, the neural familiarity mechanisms occurring during silence are similar between familiar songs and songs passively learned for two weeks, showing the efficiency of passive learning processes in encoding musical information [292].

To explain these neural responses in the absence of sound, it has been shown that silence inserted in familiar music conveys the sensation of hearing a song internally without any audible sound, which is referred to as auditory imagery. Several studies report similar spatial and temporal processes during the musical perception and imagery of a song [298, 313–315, 317]. The mental perception of a familiar piece of music occurs with a high efficiency, and not only concerns the lyrics but also the tempo [285, 287], pitch [288, 289] and timbre [290].

The present investigation sought to determine whether aesthetic musical judgement was also present during silence in music. When we like listening to Mozart, do we still like Mozart during the silences in its music? Music is a complex stimulus, and its affective dimension is an integral part of the artistic and aesthetic experience, and the hedonic appreciation of music, referred to as music liking, involves activation of areas responsible for the feeling of pleasure [318]. The hypothesis of a preservation of liking in silence is based on the important influence of familiarity on liking: becoming more familiar with a particular piece of music increases the subject's liking ratings for it [297]. Familiarity and liking being two high level information processes often co-occurring when listening

to music [319], they may as well co-occur during silences inserted in music. A second argument supporting that liking may be maintained during silence is based on the fact that emotions are preserved during auditory imagery [291].

The neural structures responsible for the mechanisms of liking have only been explored while listening to music. Neuroimaging methods have linked the activation of several cerebral areas to pleasure states when listening to music: the dorsal and ventral striatum, right caudate nucleus, ventral putamen and nucleus accumbens in relation to dopamine release in the reward system [320]. A functional architecture including anterior cingulate cortex, amygdala, insula, orbitofrontal cortex and nucleus accumbens as areas providing evaluative responses such as the reward value of the stimulus has also been suggested [308, 321]. Other structures such as the inferior frontal gyrus (IFG) and the superior temporal gyrus (STG) have been found to be specifically activated during liking/disliking processes [297, 322].

A few recent studies have described the use of EEG in listening situations to investigate the temporal dynamics of liking. Partially discordant results were found in terms of brain rhythms among these studies. When comparing consonant to dissonant music, a positive correlation between the synchronization in the mid-frontal theta band and the liking score emerged, especially towards the end of the excerpts [323]. Two other studies [324, 325] reported that beta and gamma bands were the best features to separate pleasant from unpleasant music, followed by alpha and theta bands. In a recent study a biomarker of music liking has been defined, relying on brain activations recorded over the left prefrontal cortex and focusing on the functional coupling between high-beta and low-gamma oscillations [326]. The discrepancies observed in all EEG studies may be attributed to the interference of the physical processing of sounds that impair the detection of the electrophysiological activity related to liking. Should liking processes be present during silence, we aimed to investigate whether the spatiotemporal dynamics would be similar to those observed in listening situation. To study these dynamics, we chose to use high-density EEG, whose spatial resolution allows localizing the neural structures involved in liking processes and whose high temporal resolution is adapted to the time-frequency analysis of these events. No predicted directions were given for time frequency analysis because of the discordant results obtained in previous EEG studies. Because of the high-level information that has already been shown to be preserved during silence, specific EEG liking responses were thought to emerge, with neural activations at least partially similar to those observed during listening situations.

4.3.3/ METHODS

The present study is a complementary analysis of the data from the study of [292] whose primary objective was to explore the onset of the memory dynamics that lead to the automatic completion of missing lyrics during spontaneous musical imagery.

4.3.3.1/ PARTICIPANTS

Twenty volunteers (7 male, 13 female, mean age = 32 ± 5 years old) participated in the experiment. All participants were right handed, non-musicians (i.e. had not received any formal musical training or had done so for less than 6 years), with no neurologic or psychiatric diseases and normal audition. Written consent was obtained from participants

prior to taking part in the study, and all received a monetary compensation of 50€. Research methods were approved by the Committee for the Protection of Human Subjects at the Clinical Investigation Center of Besançon (n°14/458).

4.3.3.2/ EXPERIMENTAL DESIGN

Five familiar songs chosen by the participant and five unfamiliar songs selected by the experimenter among original soundtracks recorded by independent or relatively unknown singers were used. Familiar and unfamiliar songs were paired to belong to the same musical genre, sang by a singer of the same sex. All songs were in participants' native language. Loudness levels of all songs were equalized with the Replay Gain algorithm.

The first EEG recording was performed immediately after song selection. All songs were randomly played over headphones and repeated twice. Each Song was 2 min in length and twenty portions were replaced with two-second silent intervals randomly embedded between 10 and 100 seconds after song onset and with a distance ranging from 2.5 to 3.5 seconds between gaps. This resulted in 200 trials within familiar songs and within unknown songs. Silence and song presentation were accomplished using the E-Prime software (Psychology Software Tools Inc. Sharpsburg, PA). Participants were instructed to close their eyes and remain still while passively listened to the song. They were requested to rate each song using a visual analog scale (0: unfamiliar ; 10: unfamiliar). The knowledge of each song could thus be checked to ensure that subjects did not know any of the unfamiliar songs. After the EEG recording, subjects were required to rate whether they were mentally completing the gaps during the experiment.

Then there was a two-week learning period during which each subject was inquired to listen regularly to the unfamiliar songs, and to note the frequency of song listened to per day in a self-monitoring logbook. The participants listened to each song for an average of 33.89 times (Average = 2.39 listening per day). All subjects listened to each of the five unfamiliar songs the same amount of time.

The second EEG recording was scheduled after the two-week learning period. The recording had exactly the same design as for the first EEG, at the exception of the order of the songs that was randomly selected. At the end of the recording, subjects were requested to rate subjective liking/disliking (from 0: disliked to 10: liked a lot), emotions (from 0: very sad to 10: very joyful) and familiarity (from 0: unfamiliar to 10: very familiar) for each song with a visual analog scale.

4.3.3.3/ DATA ACQUISITION AND ANALYSIS

In this study, we focused on the liking of newly-learned songs, i.e. unknown songs that were learned over the time course of two weeks, because they were controlled to be listened to the same amount of time. Among the five newly-learned songs listened to by participants, the two more preferred and the two least preferred songs were selected. Only the neural responses from the second EEG were considered.

Subjects were seated in a comfortable chair in a dark quiet testing room. EEG signals were recorded using a 256-channel Geodesic Sensor Net (Electrical Geodesics Inc. ; EGI, Eugene, OR). Data were continuously recorded with a high-pass filter at 1 Hz, and a sampling rate at 1000 Hz. Epochs of 2500 ms were extracted from the raw data beginning

600 ms pre-silence onset and ending 1900 ms post-silence onset separately for liked and disliked songs. A total of 80 trials was extracted for the more preferred and the least preferred songs. A band-pass filter between 1 and 30 Hz and a notch of 50 Hz were applied to remove unwanted frequency components. A baseline correction was applied for each trial before silence onset (-600 to 0 ms). Periods with visually detectable artifacts were removed from the analysis. Individual channels with artifacts were interpolated using a 3-dimensional spine algorithm (maximum: 7% interpolated electrodes per subject).

Nine Regions of Interest (ROIs) were selected: left anterior, central anterior, right anterior, left median, central median, right median, left posterior, central posterior and right posterior (Figure 4.2).

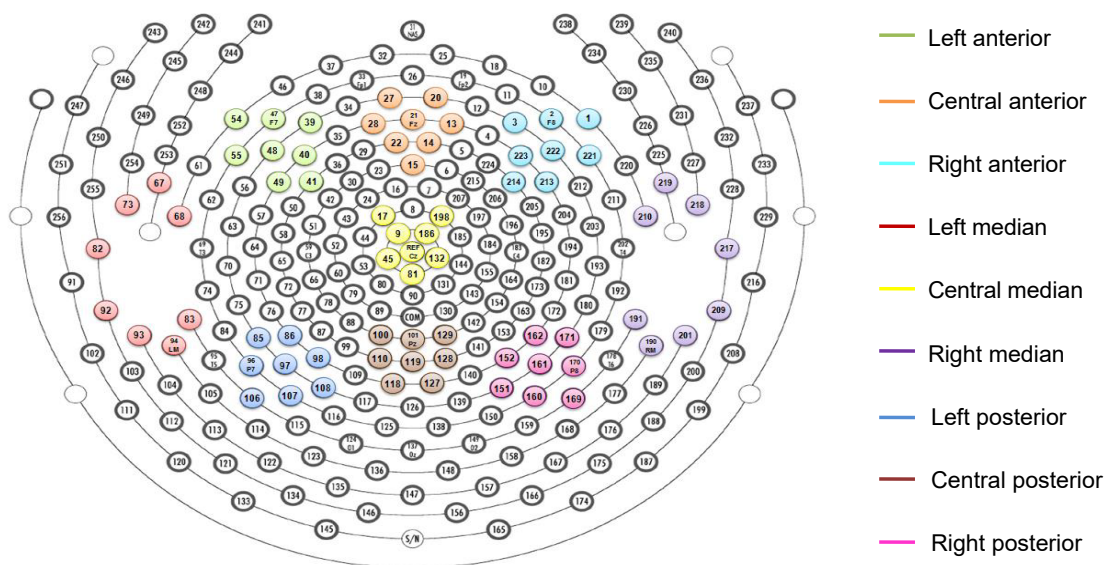


FIGURE 4.2 – The nine regions of interest. A: Anterior, M: Median, P: Posterior ; L: Left, C: Central, R: Right.

Time-frequency analysis

To evaluate the long-term signal synchronization in the frequency domain during the entire silent period, time-frequency analysis of silent was carried out in the whole time period on each ROI for liked and disliked songs. Spectral power from -50 ms pre-stimulus to 1950 ms post stimulus (in 10 ms steps). The length of the sliding time window was frequency dependent with 3 cycles at the minimum frequency of 3 Hz and 6 cycles at the maximum frequency of 30 Hz. Fifty log-spaced frequencies from 3 Hz to 50 Hz were calculated. We considered as theta the rhythmic activity occurring between 4 and 8 Hz, alpha the activity between 8.5 and 12 Hz, and beta the activity above 12 Hz and below 30 Hz. Time–frequency points were considered significant between liked and disliked songs by a non parametric permutation-based statistical method [327]. A total of 1000 permutations were used at each frequency to assess significance at $p < 0.01$. To increase statistical power, only clusters of significant responses on several frequencies and several time frames were considered as significant.

Source localization

Based on the results of the time-frequency analysis, source localization was performed on the frequency bands showing significant differences between liked and disliked songs. Source localization was estimated using a distributed linear inversed solution based on a local auto-regressive average (LAURA) model [328]. Activations obtained by source imaging were compared for each condition and thus identified diverse regional brain activities. Voxelwise two-tailed t-tests were calculated on the mean LAURA source estimates for liked and disliked songs. Differences in the current source density at $p < 0.001$ are reported.

4.3.4/ RESULTS

4.3.4.1/ BEHAVIORAL ANALYSIS

After two weeks of passive learning, the average liking rating was 5.3 ± 2.4 (range 0-10). When separating the songs into the two most liked and the most disliked for each subject, a 4 point difference was noted on average (Table 4.1). Since we selected the most and less liked music for each subject, a highly significant difference was observed between liking scores of liked and disliked songs ($t(19)=9.55$, $p < 0.001$). The comparison of emotions conveyed by liked and disliked songs revealed no significant difference ($t(19)=0.81$, $p=0.4$). Although all songs were listened to for the same amount of time, a significantly increased familiarity for a liked song was observed ($t(19)=4.79$; $p < 0.001$).

	All songs	Liked songs	Disliked songs	p value
Liking	5.4 ± 2.4	7.3 ± 2.1	3.3 ± 1.7	<0.000001
Familiarity	7.4 ± 1.6	7.8 ± 1.4	7.0 ± 1.7	<0.001
Emotions	4.8 ± 2.4	4.5 ± 2.2	5.0 ± 1.6	n.s.

Tableau 4.1 – Musical appreciation, familiarity, and emotions for all 5 songs, as well as for the most liked and disliked songs (mean followed by standard deviation). P value shows the difference of scores between liked and disliked songs.

4.3.4.2/ TIME FREQUENCY RESPONSES

Time frequency analysis revealed three distinct time periods of significant differences between liked and disliked songs: the first between 400 and 900 ms, the second between 1000 and 1300 ms, and the third between 1400 and 1900 ms (Figure 4.3). Differences were mostly significant in the theta and alpha bands, with a desynchronization of theta and a synchronization of alpha for liked songs. No differences were observed above 20 Hz.

Source localization was performed on the three time periods of difference. During the first time period, a significantly stronger theta activity was observed in the left anterior and central posterior ROIs (LA and CP in Figure 4.3). Statistics performed on source localization could not reveal the source of the differences between both conditions, but the right inferior frontal gyrus (IFG) was differentially activated at a lower significance level of $p < 0.01$ (Figure 4.4). In the second time period, a difference in theta activity was observed in the right median ROI (RM in Figure 4.3). The main difference originated from a stronger activation of the right superior temporal gyrus (STG) for disliked songs. In the last time period, a difference in alpha activity was observed in the central anterior and the

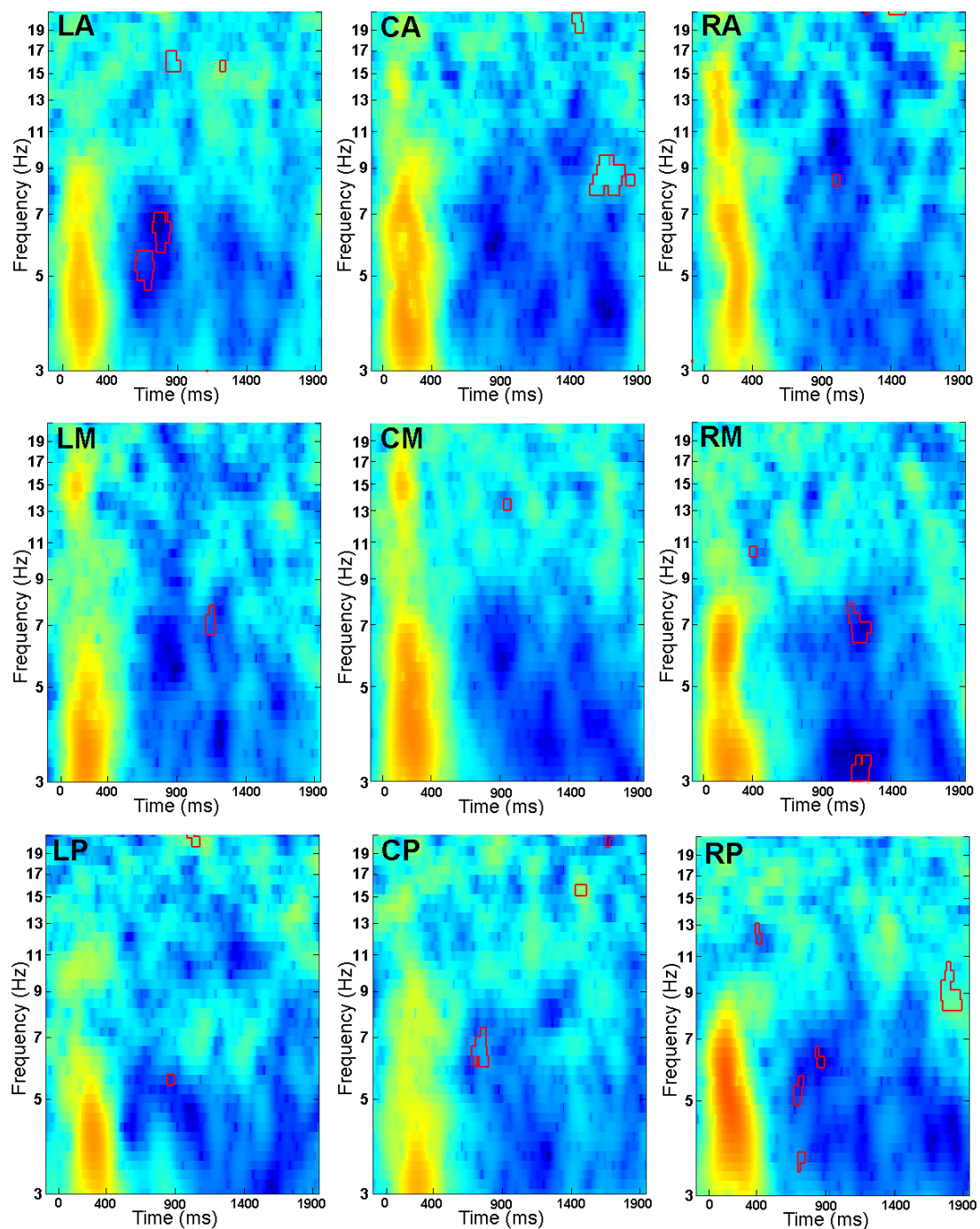


FIGURE 4.3 – Time frequency analysis performed over the nine ROIs during silence inserted into liked songs. Neural synchronizations (yellow) and desynchronizations (blue) are represented at different frequency bands. Areas of significant differences ($p < 0.01$) between liked and disliked songs are surrounded in red.

right posterior ROIs (CA and RP in Figure 4.3). Source localization showed this difference originated from the left Middle Temporal Gyrus and the right Insula.

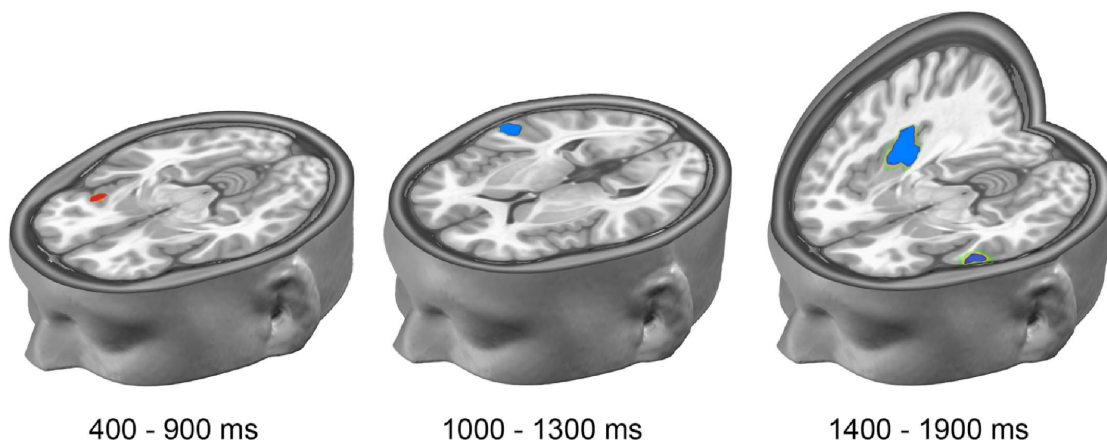


FIGURE 4.4 – Brain areas showing significant differences in activation between liked and disliked songs. Left: source of difference in theta activity in the first time period. Middle: source of difference in theta activity in the second time period ($p < 0.001$). Right: source of difference in alpha activity in the last time period ($p < 0.001$). Areas more strongly activated for liked songs are represented in red, and in blue for disliked songs.

4.3.5/ DISCUSSION

An original aspect of the present study was to explore the neural mechanisms of liking during silence, thus offering the possibility of studying its emergence while removing any acoustical interference. To study music liking in a naturalistic condition, we voluntarily chose songs instead of using artificial excerpts of music and asked the participants to simply listen to these songs for two weeks, as if they were hearing a new song on the radio, for example. Moreover, we preferred using songs as material over using instrumental music because we were interested in the integrated memory representation for text and melody in songs [329, 330]. In addition, the songs were not chosen according to the emotions they raised. Although studies on liking usually focused on happy musical extracts, since it is this type of music which most frequently leads to the highest liking scores [331, 332]. It has been shown that sad music can also be highly appreciated by listeners [333]. We thus chose the most appreciated and least appreciated songs regardless of their emotional valence. Moreover, since we have averaged over many different songs, based on preference, we thereby canceling out differences related to the stimulus.

Our results show that, although participants were instructed to simply listen to the songs in the experimental tasks, involuntary aesthetic judgements occurred during silences, which is similar to familiar judgements that have already been noted in identical conditions [284, 292].

4.3.5.1/ ELECTROPHYSIOLOGICAL CORRELATES OF LIKING

The subjective aesthetic appreciation of a song was reflected by responses in theta and alpha frequency bands in three time periods included in silent intervals. First responses, frontal theta oscillations, have not only been reported in subjects listening to pleasant musical pieces [323], but also in subjects performing a broad range of cognitive tasks requiring attention or short term memory [334–336]. Here, it is likely that theta activity observed in the two first time periods reflects both attentional and emotional mechanisms

that could be part of the same process since an emotionally significant stimulus automatically attracts attention [337]. It is, therefore, conceivable that participants were more attentive during silences occurring in pleasant songs, waiting for the music to resume so they could enjoy it again.

Variations in parieto-occipital alpha oscillations observed for liked and disliked songs could be related to auditory attention mechanisms, and it is thus possible that liked music attracted more attention than disliked songs, but this hardly explains why it occurred in such a late period. A suggestion could be that participants were eager for the pleasant songs to start again, and were attentively waiting for the end of silence gaps. However, alternative explanations could explain the increase in alpha oscillations. For example, it has also been suggested that alpha synchronization reflects an active process of inhibition of task-irrelevant areas [338]. As a consequence, variations in alpha oscillations may also be related to increased relaxation consecutive to increased enjoyment and fewer cognitive resources allocated to liked songs [339].

It should be noted that we did not find evidence that oscillations above 20 Hz were involved in music liking, which contradicts some previous studies investigating music liking through EEG signals [324–326]. In addition to the large methodological differences between these studies and ours, the main distinction lies in that liking was measured at the same time as music was perceived. In that respect, it is well known that frequencies above 20 Hz fluctuate during musical beat processing [340–342], a parameter especially difficult to control in naturalistic conditions. Studying liking during silence may have reduced the interference between beat processing and liking, even though some studies have shown that tempo was well preserved during auditory imagery [285, 287]. Of course, it is also possible that the processing of liking differs during silence and music perception and that aesthetic information is not encoded the same way in both situations. However, the difference in theta oscillations observed in our study, also found when listening to consonant and dissonant music [323], may be more a signature of music liking.

4.3.5.2/ NEURAL SOURCES OF EEG ACTIVITY

Interestingly, sources of theta activity observed during silence were found to originate from the IFG and the STG. These results corroborate previous fMRI studies that also reported a stronger activation of the IFG for pleasant music and of the STG for unpleasant music in listening situations [297, 322]. The IFG and STG are two structures known to work together to process various aspects of music [308]. They are activated for musical memory retrieval [298, 330, 343] and participate in the formation of a musical image [344]. The IFG is notably involved in the long-term musical memory retrieval [345], which is essential for making accurate predictions about music. The auditory cortex in the superior temporal region is a hub for liked, disliked and favorite music [346]. Our spatiotemporal analysis of subjective liking/disliking during silence revealed that the IFG and the STG do not work simultaneously but successively, increases in liking inducing first a stronger activation of the IFG and then a reduced response of the STG. IFG being known to have an inhibitory role on a number of brain systems and related behaviors [347], this structure may inhibit the STG for processing information related to liked songs.

In the last period of silence, the differential activation of the insula suggests that hedonic processes are involved. The insula contributes to the conscious experience of emotion and its activation covaries with subjective feelings, which not only reflect physical stimulus

intensity but also cognitive factors such as prediction [348].

There were surprisingly few cerebral areas activated for the aesthetic judgement of music in our study, which differs from the large number of cerebral activations observed in fMRI studies [297, 322, 346]. Since EEG surface signals are predominantly produced by cortical structures, source localization methods have difficulties to localize deep brain structures. This may explain why structures essential in liking processes such as the limbic system structures and the hippocampus [346], could not be detected in our study. Moreover, liking scores were not greatly different between liked and disliked songs, only 4 points on average, whereas some studies have been investigating participants' favorite songs [346] or compared dissonant and consonant music [323]. Finally, a possibility is that aesthetic judgement is altered during silence compared to during music perception, leading to a partially activated network of neural activations.

4.3.5.3/ RELATIONSHIP BETWEEN LIKING, EXPOSURE, AND FAMILIARITY

Since the liking for unfamiliar music is known to increase as a function of exposure time [349, 350], we wanted to avoid any confounding effect by assuring that participants listened to liked and disliked songs the same number of times. Despite this control, we found that liked songs were rated as more familiar than disliked songs, confirming that listeners do not only like music that they remember but also remember music that they like [319]. This strong relationship between preference and familiarity is confirmed by source localization. The IFG and STG are two structures involved both in memory processes and in liking. Moreover, emotion-related limbic and paralimbic regions as well as the reward circuitry are known to be significantly more active for familiar relative to unfamiliar music [297]. If musical preference depends on predictive processes based on our knowledge of a song, it is then obvious that familiarity plays an important role. Further research is necessary to clarify this point, although it would be extremely difficult, if not impossible, to measure musical preference without any interference from familiarity and exposure.

4.3.6/ CONCLUSION

Results from the present study show that we like listening to a song, we still like this song during the silences inserted inside. During silence, neural responses specific to liking emerged under the form of oscillations in theta and alpha frequency bands originating from the IFG and STG. These brain structures having previously been shown to be activated when listening to liked music, similar neural networks may be involved in silence and music. However, all liking processes may not present because there were less cerebral areas activated for the aesthetic judgement of music in silence. The experimental design used here nonetheless offer the opportunity to study the emergence of the spatio-temporal dynamics of liking without physical interference of music.

4.4/ COMPLÉMENT À L'ARTICLE PRÉCÉDENT : ANALYSE DES ERPS

4.4.1/ MATERIAL AND METHODS: EVENT RELATED POTENTIAL ANALYSIS

ERP waveforms were averaged in each ROI and TW and mean amplitudes were submitted to repeated measures analysis of variance (ANOVA) using the following factors: Liking (liked/disliked), Gradient (anterior/median/posterior) and Laterality (left/central/right).

4.4.2/ RESULTS

The visual analysis revealed distinct ERPs in the four TWs of interest for both liked and disliked songs. Statistical analysis revealed an interaction between Liking and Gradient in TW2 ($F(2,38)=3.85$, $p<0.05$) and TW3 ($F(2,38)=4.96$, $p<0.05$). Post-hoc analysis (Fisher's LSD corrected) revealed that musical preference resulted in a difference in amplitude in the anterior region for both TWs (TW2: $p<0.05$; TW3: $p<0.01$) (Figure 4.5). In TW2 and TW3, statistics performed on source localization revealed that this difference mainly originated from a stronger activity in the middle and right IFG for liked songs. Moreover, the STG was also more strongly activated in TW2 for liked songs. No significant differences were found between liked and disliked songs in TW1 and TW4.

4.4.3/ DISCUSSION

Inserting silent intervals resulted in the generation of an early error response occurring at the same time as anticipation and expectations from music were processed. The error response corresponded to a violation of the listener's prediction and was modulated by musical preference. Two TWs were specifically altered. The latency and topography of the first one are similar to a mismatch negativity (MMN) component, a pre-attentive error response that reflects a deviation from expected musical patterns [311, 312]. The duration of the MMN was shorter for disliked songs, which could be attributed to a reduced prediction violation. The positive distribution of TW3 was similar to a P3a, an ERP frequently evoked after the MMN and that can also be interpreted as an error response. Because the P3a reflects a process of selecting information within working memory, the cognitive load may be reduced when having to spontaneously complete liked songs. For both the MMN and the P3a, the neural source of difference was localized in the right IFG, a structure known to be responsible for the generation of the MMN [351] and that was also involved in differences in theta oscillations in the first time period. The IFG has multiple roles in music processing, and we show here that it is used both in retrieving stored musical information for a preference judgment and to indicate a violation of new musical expectancies.

4.5/ CONCLUSION

Cette étude a permis de mettre en évidence une activité cérébrale spécifique liée à l'appréciation musicale, qui se traduit par des réponses spécifiques à court terme (MMN et P3a), et par une activité à long terme (synchronisation/désynchronisation de l'activité

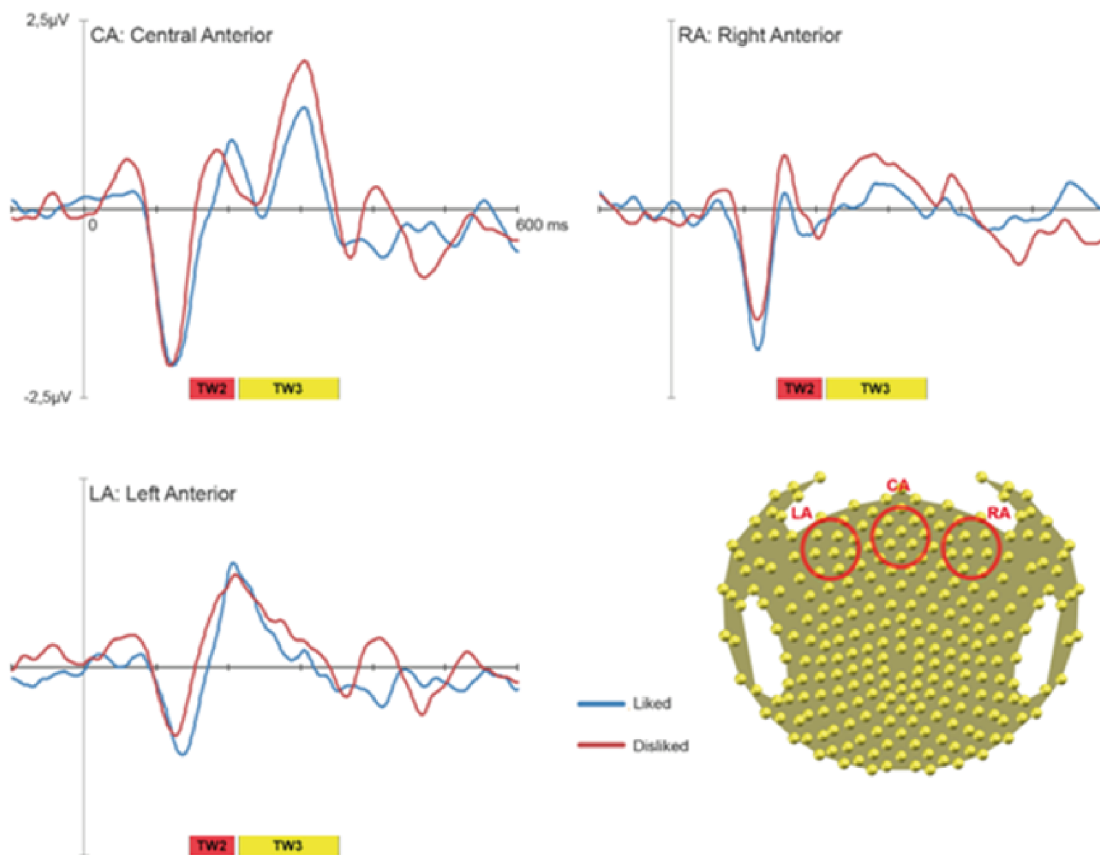


FIGURE 4.5 – ERPs for liked and disliked songs in the three anterior ROIs. A significant effect of gradient was obtained, with neural differences in musical preference on anterior ROIs for both TW2 and TW3.

rythmique Thêta et Alpha). Ce protocole nous permet de distinguer de manière automatique les chansons appréciées de celles qui ne le sont pas, ouvrant des perspectives concernant l'amélioration de l'encodage et du rappel par neurofeedback pour moduler ces rythmes cérébraux, ou en stimulant artificiellement les structures cérébrales concernées.

SVM APPLIQUÉ AU PROJET RIHANNA

Dans ce chapitre, nous appliquons la classification SVM aux données issues du projet Rihanna. Nous souhaitons examiner la possibilité de discriminer automatiquement l'activité cérébrale issue de l'imagerie musicale sur chansons familières et inconnues, de façon individuelle.

Cette étude fera l'objet d'un article, actuellement en préparation sous le titre de travail *Machine learning individual differences in the detection of electrophysiological responses of musical familiarity*.

5.1/ RÉSUMÉ DE L'ARTICLE

5.1.1/ INTRODUCTION

La prévalence de la maladie d'Alzheimer et de la dépression, qui touchent fortement la population âgée, est en constante augmentation, mais des symptômes communs, en particulier au niveau des fonctions mnésiques, rendent difficile leur discrimination. Un protocole combinant musique et imagerie cérébrale a été développé par Gabriel *et al.* (2016) [292] pour estimer objectivement les capacités mnésiques par EEG-HR. Le principe est d'écouter passivement des chansons familières et inconnues du patient, dans lesquelles sont insérées de courtes périodes de silence. Si la chanson est connue, le cerveau tente automatiquement de combler le silence, ce qui n'est pas le cas si la chanson est inconnue. Toutefois, si ce protocole donne des différences significatives sur l'analyse de groupe, sa transposition à l'échelle individuelle, nécessaire pour le diagnostic, est plus délicate. Nous nous sommes donc dirigés vers le SVM, une méthode d'apprentissage machine récente, pour discriminer automatiquement les essais des deux conditions.

5.1.2/ MATÉRIEL ET MÉTHODES

L'activité cérébrale de vingt volontaires sains a été enregistrée par EEG-HR pendant l'écoute de 20 chansons familières ou inconnues, chacune entrecoupée de 21 périodes de silence. Nous proposons un pipeline consistant en une étape préliminaire de prétraitement des données à l'aide de la méthode xDAWN, puis en l'exploration de la meilleure façon d'extraire les informations pertinentes des enregistrements EEG tant dans le domaine temporel que fréquentiel.

5.1.3/ RÉSULTATS ET DISCUSSION

Les résultats montrent des différences significatives de la performance de classification en fonction de la méthode d'extraction des caractéristiques du signal utilisée. Une différence de performance est également observée d'un sujet à l'autre, confirmant la présence d'une forte variabilité inter-individuelle. Enfin, notre étude souligne la nécessité d'optimiser les hyperparamètres du SVM pour chaque sujet d'intérêt.

5.2/ MACHINE LEARNING INDIVIDUAL DIFFERENCES IN THE DETECTION OF ELECTROPHYSIOLOGICAL RESPONSES OF MUSICAL FAMILIARITY

Joucla Coralie¹, Gabriel Damien^{1, 2}, Ortega Juan-Pablo^{3, 4}, Haffen Emmanuel^{1, 2, 5} and Grigoryeva Lyudmila⁶

¹ Neurosciences intégratives et cliniques EA 481, Université de Franche-Comté, F-25000 Besançon, France

² INSERM CIC 1431, CHRU Besançon, F-25000 Besançon, France

³ Centre National de la Recherche Scientifique (CNRS)

⁴ Faculty of Mathematics and Statistics, University of St. Gallen, Switzerland

⁵ Service de Psychiatrie de l'adulte, CHRU, F-25000 Besançon, France

⁶ Department of Mathematics and Statistics, Universität Konstanz, Germany

5.2.1/ ABSTRACT

Brain electrical activity can be observed by simply putting on ElectroEncephaloGraphic (EEG) electrodes on the scalp. Identification of this activity allows to single out brain areas and temporal decay triggered by a specific stimulus. However, each and every though a subject has in the mind, independently of the experimental tasks, elicits electrical activations within different parts of the brain. These activities represent a noise often more important than the signal we ought to detect. In conventional studies, this problem is overcome by averaging a great number of trials across subjects. Random activity disappears while activity of interest emerges. However, while these methods allows for fundamental studies, they are difficult to transpose to a clinical approach where a single subject examination is needed. In recent years, machine learning methods based on EEG measurements rose in popularity in medicine fields. Machine learning methods have the ability to discriminate patterns in data that are hardly or even not distinguishable with conventional techniques, with the potential outcome of moving from group-level statistical results to personalized diagnosis in neuro-psychiatric pathologies.

However, applying machine learning on EEG data can be challenging, because of the low signal-to-noise ratio and of a high inter-individual variability. Thus, we propose a pipeline using the xDAWN method for data cleaning, as well as an exploration of the best way to extract relevant information from EEG recordings. Support Vector Machine (SVM) being the most commonly used machine learning technique with EEG, we chose to apply it on

our data to build individual classification models. Finally, SVM comes with several hyperparameters that are often not taken into account in literature. Results show a significant difference in classification accuracy and pValue depending on the chosen features and hyper parameters. Moreover, the optimum features and parameters may vary subject to subject. Our study highlights the necessity for these hyperparameters and features to be tuned for each test subject.

5.2.2/ INTRODUCTION

Diagnostic difficulties are frequent in medicine, regardless of the medical field concerned. This difficulty is exacerbated in psychiatry and neurology, due to the lack of robust diagnostic criteria such as pathognomonic biological or physiological markers. In addition, borderline, progressive and comorbid pathologies are often observed [2]. The implementation of effective therapeutic strategies requires rapid and reliable diagnosis, as there is often insufficient time to limit the evolutionary risks, as is the case for Alzheimer's disease and depression. The prevalence of these two pathologies, which strongly affect the elderly population, is constantly increasing, but common symptoms, particularly in terms of memory functions, make it difficult to discriminate between them [270]. Indeed, the memorisation difficulties of Alzheimer's disease and depression obey totally different mechanisms. Alzheimer's disease is characterised by an impairment of the hippocampal pathway, *i.e.* a deficit in recording information [3, 4]. In subjects suffering from depression, and more particularly in subjects suffering from pseudo-dementia, the difficulties are related to the retrieval of information but not to its storage [25].

To overcome these drawbacks, an original protocol combining music and brain imagery has been developed at the Clinical Investigation Centre of Besançon (France) to estimate memory impairment without the active participation of the patient [292]. The principle is to passively listen to songs familiar or unknown to the subject, in which short periods of silence are inserted. If the song is known, the brain automatically tries to fill the silence, while it is not the case if the song is unknown. Thus, by measuring the brain activity when listening to these songs and especially to the silences, it would be possible to objectively measure the memory capacities. This identification of specific neuromarkers would be potentially more effective and safer than the behavioural assessments used in clinical examinations, since the same behaviour can result from various cognitive dysfunctions.

To evaluate the effectiveness of the protocol, twenty healthy volunteers underwent an initial high-resolution electroencephalography (HR-EEG) examination during which they listened to 20 familiar or unknown songs, each interspersed with 21 periods of silence. The effects of familiarity appeared in two different forms during the beginning of the silence: firstly, in the form of lower amplitude evoked potentials for unknown songs, reflecting an inability to make accurate predictions about the song. And secondly, a decrease in power on the theta band of interest was observed for unknown songs throughout the duration of the silence. This second variation would reflect greater difficulty in activating areas of the brain related to memory.

However, while significant differences were reported at the group level, most of these results do not distinguish between patients at the level of a single subject. Indeed, the EEG suffers from some limitations that hinder its effective analysis and treatment. Firstly, the EEG has a low signal-to-noise ratio, as recorded brain activity is often stacked with multiple sources of environmental, physiological and activity-specific noise of similar or

greater amplitude, called “artifacts”. Although various noise filtering and reduction techniques have been developed [150, 185, 192, 236], EEG researchers still often rely on averaging tests to extract actual brain activity from the signal. This works very well for group studies, but induces issues when trying to use it at the level of a single subject, as it involves using fewer trials. Secondly, brain activity can vary considerably from subject to subject [352]. This wide inter-subject variability is one of the main drawbacks of using the EEG in the clinical field.

In recent years, machine learning methods, especially the Support Vector Machine (SVM), have gained popularity in EEG signal processing because of their success in discriminating patterns in data that are difficult or even impossible to distinguish with conventional techniques. SVM has the potential to move from statistical results at the group level to personalised results.

We therefore examined which SVM protocols were most commonly applied to EEG recordings in order to draw lessons for our study. We found that the features used are often extracted from the frequency domain [114, 194, 210], but that the time domain could also be useful in the classification of evoked potentials [126, 127, 225]. We also found that the selection of the most relevant features was important for good classification accuracy [115, 154, 184, 191]. We therefore chose to explore different feature extraction methods, on all electrodes and on groups of electrodes based on the regions of interest that showed significant differences in our previous study [292].

Moreover, it seems that the choice of the SVM architecture and the cross-validation parameters are paramount for the classification performance. We examine in detail the mechanisms and hyperparameters of SVM that can affect classification performance, in addition to the data itself.

5.2.3/ MATERIAL AND METHODS

5.2.3.1/ PARTICIPANTS

Twenty healthy right-handed subjects (7 male, 13 female, mean 32 ± 5 y.o.) participated in this study. All reported normal hearing, with no neurologic or psychiatric disorders. Neuropsychological evaluations revealed no memory deficits.

The Committee for the Protection of Human Subjects of Besançon (France) approved this experiment under N° 14/458, and written informed consent was obtained from all participants.

5.2.3.2/ EXPERIMENTAL DESIGN

Full experimental design was first reported in Gabriel *et al.* (2016) [292], which data are used in this study. It consisted of 3 main phases: a first EEG recording, a two weeks long learning period, and a second EEG recording. Only the data from the first recording was processed for this paper, thus the methodology about the learning period and the second EEG will not be detailed here.

Prior to the first recording, participants freely chose 5 familiar songs in their native language (French). The researchers then selected 5 songs, unfamiliar to the subject, but

matched in terms of musical genre and singer's gender.

The Replay-Gain algorithm was then used to normalize the songs to the same loudness level. Two seconds silent periods were randomly inserted in place of music between 10 and 110 s after song onset, and interspaced of 2.5 to 3.5 seconds from each other. Songs were then cut to last 2 minutes and include 21 silences, and repeated twice, resulting in 210 trials (21 gaps x 5 songs x 2 repetitions) for each condition (familiar versus unfamiliar).

5.2.3.3/ DATA ACQUISITION

During the EEG recording, the songs including the silent gaps were presented in a random fashion and played through headphones. Subjects settled the volume to be at ease, passively listening to the musics, eyes closed, in a quiet experiment room. The recording lasted about 50 minutes. A break was taken every 5 songs to allow the subject to stretch and rest, and thus ensure their alertness. All EEG channels were recorded using a 256 channels Geodesic Sensor Net (Electrical Geodesics Inc; EGI, Eugene, OR), a high impedance amplifier NetAmp 300 Amplifier (Electrical Geodesics Inc.) and collected via the NetStation 4.5 software (Electrical Geodesics Inc.). A sampling frequency at 1000 Hz and a high-pass filter at 1 Hz were set. All channels were referenced to the vertex electrode (Cz).

5.2.3.4/ PREPROCESSING

Preprocessing steps were performed as follow: epochs ranging from 600 ms pre-silent gap onset to 1900 ms post-silent gap inset were extracted for each condition and participant. A second order Butterworth filter (roll-off -12 dB/oct) band-passed the signals between 1 and 30 Hz and a notch of 50 Hz eliminated the power supply artifacts. Epochs were finally baseline corrected using the average EEG voltage ranging from 600 ms to 100 ms before silence onset, for each trial and channel. Electrodes exhibiting voltages more than 100 μ V were interpolated using a 3D spine algorithm, with no more than 7% interpolated channels per participant. Channels were then recomputed to common average reference.

It appeared for 7 subjects that the stimuli markers of the last song of each recording session shifted because of desynchronization between the recording and the stimulating software, leading to class imbalance. Thus, we carried out an undersampling method as follows: as trials from entire songs were missing, trials from the last song from the majority class to be heard were chopped off to balance the two classes.

The xDAWN spatial filter purpose is to enhance the Event Related Potential so that it stands out of the noise, by maximizing the signal to signal-plus-noise ratio in single trial studies. We wrote the algorithm following the Rivet *et al.* (2009) [118] recommandations, using the MathWorks® software Matlab R2016b®.

5.2.3.5/ FEATURES EXTRACTION

Features were extracted in several ways to explore which one would be the most representative of the two classes, and thus the most efficient for the classification (Tableau 5.1).

- To deal with the huge number of channels (257) while keeping the time course of the ERP, we chose to compute the Global Field Power (GFP) [72], a well-known spatial filter in EEG field. The GFP can be understood as the standard deviation of the mean of all voltage values across averaged referenced electrodes at each time point, for each trial.
- The log-power values of each of the coefficients obtained from the Fast Fourier Transform (FFT) were computed for the signals coming from each of the selected channels between 1 and 30 Hz with a resolution of 0.49 Hz.
- After the whole FFT was computed, the log-power values were averaged over the four following frequency ranges: 4-8 Hz, 8.5-12 Hz, 12.5-18 Hz, and 18.5-30 Hz.
- In Gabriel *et al.* (2016) [292], electrodes were apportioned into 9 regions of interest (ROI), each one of them composed of 8 electrodes: left anterior, central anterior, right anterior, left median, central median, right median, left posterior, central posterior and right posterior (Figure 5.1). We chose to apply the feature extraction methods not only to the 257 available channels, but also to each one of the ROI, as the following: GFP, mean, whole FFT and 4 frequency bands FFT on each electrode individually and on the mean of each ROI.

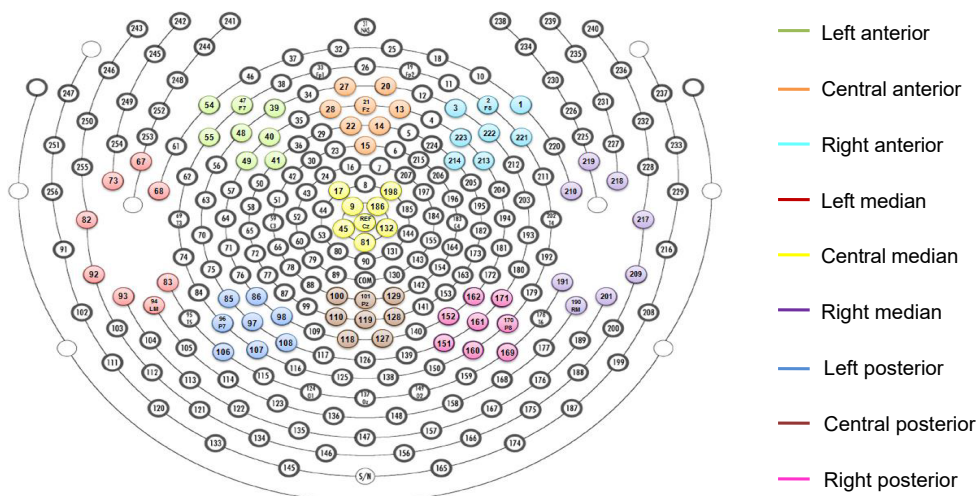


FIGURE 5.1 – Regions of Interest (ROIs) used for classification, based on Gabriel *et al.* (2016). [292]

5.2.3.6/ SVM CLASSIFICATION

Among classifiers, SVM have been widely implemented for EEG data thanks to their robustness with non-stationary and high dimensional data. A SVM uses a linear decision boundary to classify new samples. There are typically many lines that separate the data, and training the SVM corresponds to choose the decision boundary, *i.e.* the hyperplane presenting the largest margin between the two conditions.

The Matlab function `svmtrain` allows for different methods to solve the SVM regression problem. The quadratic programming method was selected for both linear and RBF kernels, for it is higher in precision when solving the SVM regression problem.

<u>Features</u>	<u>Abbreviation</u>	<u>Type</u>	<u>Features vector size</u>
All 257 electrodes			
GFP computed on 257 channels	allGFP	Temporal	1900 (1 x 1900 ms)
FFT computed on 4 frequency bands on each one of the 257 channels	all4bFFT	Frequency	1028 (4 x 257 chans)
Whole FFT computed on each one of the 257 channels	allWhFFT	Frequency	15163 (59 x 257 chans)
9 ROIs of 8 channels each			
GFP computed on 8 channels	roiGFP	Temporal	1900 (1 x 1900 ms)
Mean computed on 8 channels	roiMean	Temporal	1900 (1 x 1900 ms)
FFT computed on 4 frequency bands on each one of the 8 channels	roi4bFFT	Frequency	32 (4 x 8 chans)
FFT computed on 4 frequency bands on the mean of 8 channels	roiMean4bFFT	Frequency	4 (4 x 1 chans)
Whole FFT computed on each one of the 8 channels	roiWhFFT	Frequency	472 (59 x 8 chans)
Whole FFT computed on the mean of 8 channels	roiMeanWhFFT	Frequency	59 (59 x 1 chans)

Tableau 5.1 – Summary of feature extraction methods: method, name abbreviation, feature type and features vector size.

Kernels and gridsearch

Kernel methods map the data into higher dimensional spaces in the hope that the data could become more easily separated in this new space. The Linear kernel for example is the simplest kernel function, which depends on the hyperparameter C . Radial Basis Function (RBF) allows to pick out circles (or hyperspheres) – in contrast with the Linear kernel, which allows only to pick out lines (or hyperplanes), and depend on two hyperparameters, C and σ .

EEG data being characterized by high levels of noise and outliers, data is most of the time non-linearly separable. In that case, SVM will produce a model that allows the differentiation of many, but not all trials, called soft margin SVM.

- The regularization parameter C allows modulating the penalty associated by the classifier to these errors. The higher the C , the stricter the classifier is during training, leading to a better training accuracy but also to a very complex model that takes into account all the training points. By searching for the optimal value of C , the user is thus able to trade the training set accuracy off for better generalization ability, ensuring a better testing accuracy.
- Sigma determines the width for Gaussian distribution. For a larger sigma, the decision tends to be flexible and smooth, it tends to make wrong classification while training, but avoids the hazard of overfitting in testing set.

The choice of C and sigma being critical, and as no optimal criteria are given to set the value, we considered grid search and cross-validation to optimize the hyperparameters, *i.e.* we systematically evaluate performance of each possible combination of hyper parameters within a given range: from 10^{-7} to 10^5 for C and from 1 to 10^3 for Sigma. SVM by default sets the C and Sigma to be 1.

We will do this for each test subject individually, on each type of features extracted, on all 257 channels or on ROIs.

Crossvalidation

In machine learning, cross-validation consists in holding out a certain percentage of samples to validate the performance of the model on unseen trials. Here we use a classical approach consisting in dividing the data into 20 folds, 19 of them being used to train the model, and the spared one to validate it. This training-validation operation is repeated on each of the 20 folds, and the whole cross-validation is done for each combination of hyper parameters. The validation performance can be assessed through the mean of the chosen metric, *i.e.* accuracy and pValue over the 20 validation sets.

However, if this division of the whole dataset allows assessing the best combination of hyper parameters, one must test the best learned function on unseen data to avoid overfitting and assess the generalization ability of the model. We therefore keep 10% of the samples outside of the cross-validation loop to constitute the test set, and we chose to do 50 permutations to assess the overall performance on different test sets (Figure 5.2).

5.2.3.7/ STATISTICAL SIGNIFICANCE: PERMUTATION TESTS

As the number of tests retained is almost identical from one class to another, the hypothesis H_0 in a statistical test on the classification results is that the accuracy is 50%. We therefore used a permutation test which consists in constructing an empirical distribution of the accuracies evaluated for the different possible re-labelling of the trials, in cross-validation. In other words, we performed a cross-validation again by randomly modifying the class of tests, out of 1000 permutations. With such a test, the probability of obtaining better classification accuracy than that achieved when using the correct labels is a sign of a lack of statistical significance.

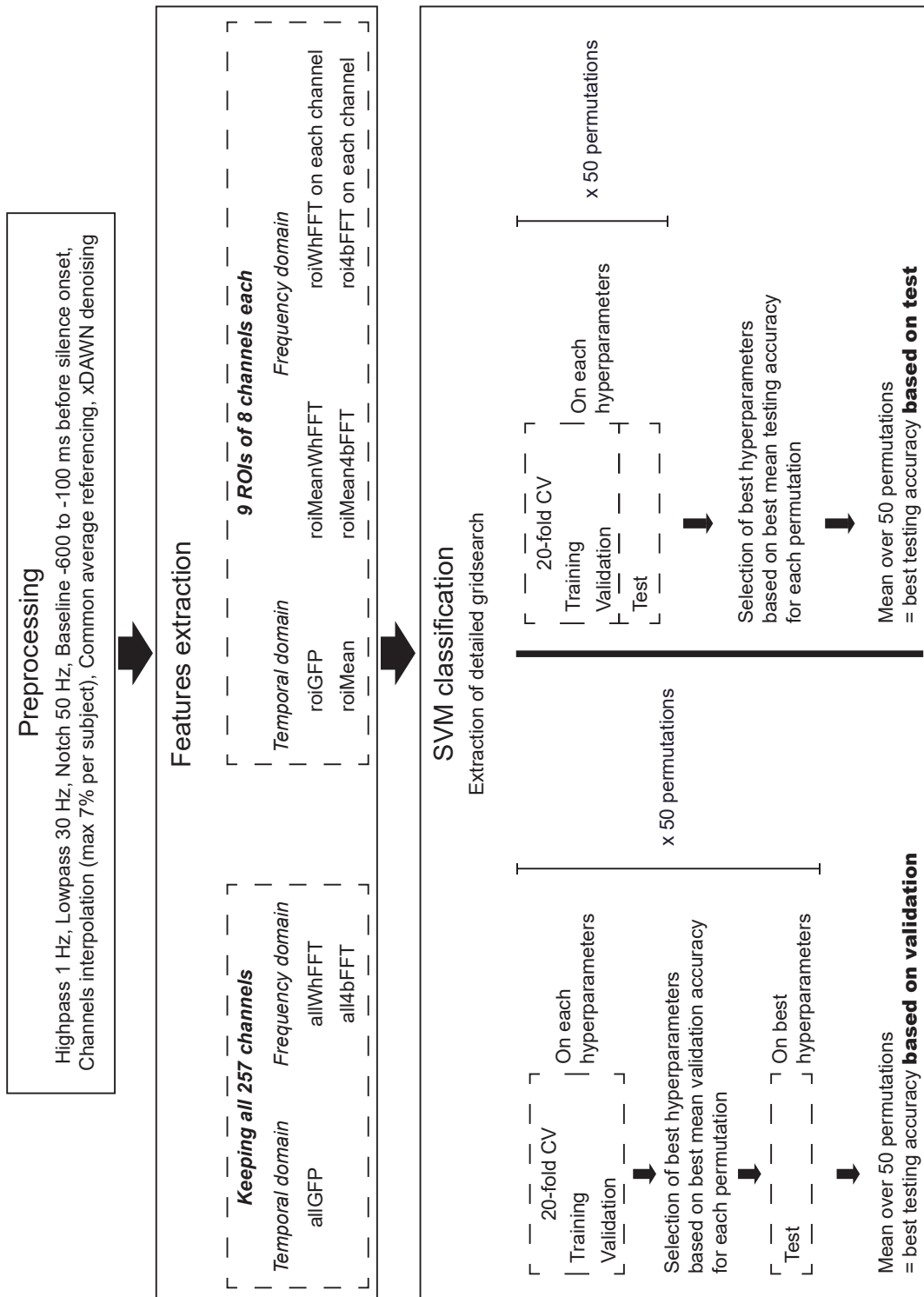


FIGURE 5.2 – Preprocessing and classification pipeline proposed for each subject individually.

5.2.4/ RESULTS

5.2.4.1/ CLASSIFICATION ACCURACIES

When looking at the results based on validation (Figure 5.2), participants achieved a mean performance accuracy of $X = 62.8\%$ ($SD = 4.5\%$; range = [55.1% - 70.6%]). Eighteen participants obtained accuracies higher than chance level, estimated to be 55.6% of correct classification accuracy for binary classification with more than 160 trials per class and $\alpha = 0.05$ [353] for both linear and RBF kernels.

For the linear kernel, mean performance accuracy was $X = 61.8\%$ ($SD = 4.8\%$; range = [54.63% - 70.6%]) (Table 5.2), and while for the RBF kernel it was $X = 62.6\%$ ($SD = 4.4\%$; range = [54.57% - 70.1%]) (Table 5.2).

Best accuracy based on validation						
Subject	Linear kernel			RBF kernel		
	Features	Accuracy	Acc.< 55,6	Features	Accuracy	Acc.< 55,6
1	all4bFFT	70,6%	1	roiMean4bFFT	68,0%	1
2	roiMean	59,8%	1	roiMean	60,8%	1
3	roiWhFFT	69,4%	1	roiWhFFT	70,1%	1
4	allGFP	63,3%	1	roiGFP	62,4%	1
5	roi4bFFT	64,0%	1	all4bFFT	64,2%	1
6	roi4bFFT	54,6%	0	roi4bFFT	55,1%	0
7	roi4bFFT	61,1%	1	roi4bFFT	61,9%	1
8	roi4bFFT	59,3%	1	all4bFFT	62,6%	1
9	roi4bFFT	68,1%	1	roi4bFFT	69,1%	1
10	roi4bFFT	65,0%	1	allWhFFT	65,7%	1
11	roiMean4bFFT	58,7%	1	roi4bFFT	62,3%	1
12	roiMeanWhFFT	64,2%	1	allWhFFT	63,9%	1
13	roiMean	57,2%	1	roiMean	60,3%	1
14	roiMean	58,5%	1	roiMean	59,7%	1
15	roi4bFFT	58,6%	1	roi4bFFT	59,6%	1
16	allGFP	69,0%	1	roiMean	68,9%	1
17	roi4bFFT	55,3%	0	roiMean	54,6%	0
18	roiMean4bFFT	61,4%	1	roiMean4bFFT	62,6%	1
19	roi4bFFT	61,2%	1	roiMean	63,4%	1
20	roi4bFFT	55,9%	1	roiMean	56,5%	1
	Mean	61,8%	pValue	Mean	62,6%	pValue
	STD	4,8%	< 0,05	STD	4,4%	< 0,05
	Min	54,63%	< 0,01	Min	54,57%	< 0,01
	Max	70,6%	< 0,001	Max	70,1%	< 0,001

Tableau 5.2 – Best testing accuracy based on validation for each subject and kernel. The best accuracy is determined across all ROIs and features extraction methods examined. Third and sixth columns: in blue colour, significance of the pValue obtained by permutation test. Logical indicator, if the accuracy is higher than 55.6%, flag = 1, otherwise flag = 0, according to the Müller-Putz (2008) accuracy significance test [353].

5.2.4.2/ KERNELS

For 15 subjects, RBF kernel gave accuracies superior than did linear kernel, for a mean improvement of $X = 0.8\%$ ($SD = 1.5\%$; range = [0.2% - 2.6%]) (Table 5.3).

Mean CPU times across participants to compute all classification cycles for each features extraction methods was $X = 68$ hrs 22 min ($SD = 14$ hrs 40 min; range = [40 hrs 0 min – 93 hrs 15 min]) for linear kernel and $X = 2546$ hrs 23 min ($SD = 202$ hrs 38 min; range = [2238 hrs 40 min – 2922 hrs 0 min]) for RBF kernel. CPU times for RBF kernel were 39 times longer than those obtained for linear kernel in average ($X = 39$; $SD = 10$; range = [26 - 61]) (Table 5.3).

Best kernel	Maximal accuracy	Difference of accuracy (linear - RBF)	CPU computation times (h)		Quotient (CPU time RBF / CPU time linear)
			Linear kernel	RBF kernel	
Linear	70,6%	2,6%	40:00	2421:48	61
RBF	60,8%	-1,0%	60:04	2704:57	45
RBF	70,1%	-0,6%	53:13	2382:43	45
Linear	63,3%	0,9%	70:33	2538:43	36
RBF	64,2%	-0,2%	82:35	2374:38	29
RBF	55,1%	-0,5%	82:01	2238:40	27
RBF	61,9%	-0,8%	56:44	2580:40	45
RBF	62,6%	-3,3%	52:56	2444:04	46
RBF	69,1%	-1,0%	53:53	2452:59	46
RBF	65,7%	-0,7%	75:45	2379:13	31
RBF	62,3%	-3,6%	55:56	2791:58	50
Linear	64,2%	0,3%	65:13	2681:00	41
RBF	60,3%	-3,2%	63:37	2922:00	46
RBF	59,7%	-1,2%	64:28	2878:59	45
RBF	59,6%	-1,0%	84:19	2355:10	28
Linear	69,0%	0,1%	87:32	2270:44	26
Linear	55,3%	0,7%	91:00	2663:39	29
RBF	62,6%	-1,1%	93:15	2405:34	26
RBF	63,4%	-2,2%	67:48	2762:25	41
RBF	56,5%	-0,6%	66:32	2677:45	40
Mean	62,8%	-0,8%	68:22	2546:23	39
STD	4,5%	1,5%	14:40	202:38	10
Max	70,6%	2,6%	93:15	2922:00	61
Min	55,1%	0,2%	40:00	2238:40	26

Tableau 5.3 – Best testing accuracy and kernel based on validation for each subject. Computation times ratio between kernels is given by RBF kernel CPU time divided by linear kernel CPU time.

5.2.4.3/ FEATURES EXTRACTION METHODS

We examined which features extraction methods gave the highest accuracy for each subject, kernel and for the use of all channels or ROIs (Table 5.2, Figure 5.3).

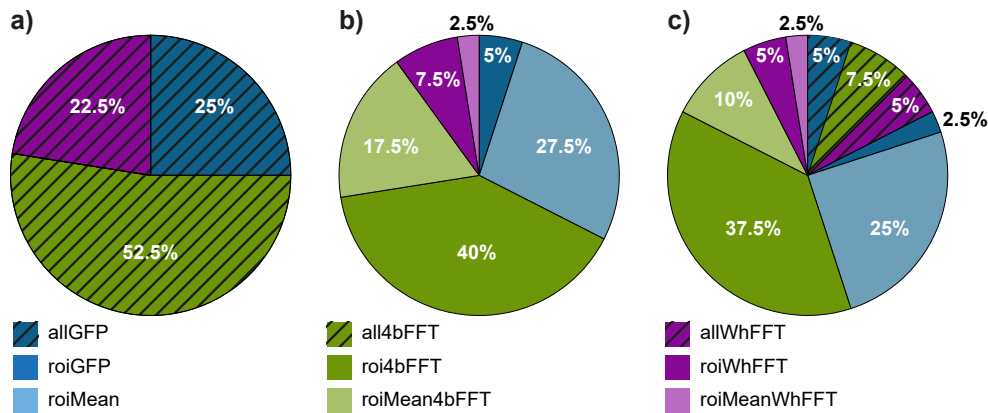


FIGURE 5.3 – Proportion of subjects for whom a particular features extraction method is best regarding classification accuracy, for both linear and RBF kernels. a) Methods using all of the 257 channels. b) Methods using ROIs. c) Best over 257 channels and ROIs.

When considering the use of all 257 available electrodes, the use of FFT on 4 frequency bands is the most effective method for 52.5% of the subjects, while the other two methods (GFP and full FFT) are more effective for 25% and 22.5% of the subjects respectively (Figure 5.3 a).

The pie chart from Figure 5.3 b shows the proportion of subjects for whom a particular method is best when using ROI (regardless of which ROI is given as the best). The majority of the subjects (57.5%) obtain the best results by using the Fourier transform on 4 frequency bands, 40% by taking the 8 ROI electrodes individually, and 17.5% by taking the average of the electrodes. 32.5% of the subjects are best represented by the features in the time domain: 27.5% for the average of the electrodes, and 5% for the GFP. Finally, only 10% of the subjects obtain the best results using the complete Fourier transform: 7.5% when taking the electrodes individually, and 2.5% when taking the average of the electrodes.

Finally, the pie chart from Figure 5.3 c shows the proportion of subjects for whom the best accuracy is obtained, depending on the features extraction method and regardless of the number of electrodes used (all electrodes or ROI). A vast majority (82.5%) of subjects have a better performance when a ROI is used. Of these, 37.5% are best represented by the Fourier transform over 4 frequency bands on the individual electrodes of a ROI and 25% are best represented by the average of the ROI electrodes in the time domain.

5.2.4.4/ REGIONS OF INTEREST

Moreover, in order to confirm which ROI and which method gives the best results in classification, we performed 3-factor ANOVAs (Frontal position, Sagittal position, Features extraction method) for each subject and for each of the 2 kernels.

Table 5.4 shows the statistical results of the 3-factor ANOVA performed for the linear

kernel. On 20 subjects, 19 reach the significance threshold ($p < 0.05$) for the “features extraction method” factor, including 18 whose pValue is lower than 0.001. The pValue is significant for 2 subjects for the “frontal position” factor and for 1 subject for the “sagittal position” factor. The interactions between sagittal and frontal position and between frontal position and type of features were significant only for 2 and 1 subjects respectively ($p < 0.05$). However, when Bonferroni-correcting these analyses, we found that only the “features extraction method” factor reached the significance threshold ($\alpha < 0.025$).

Subject	Frontal position	Sagittal position	Features extraction method	Frontal position x Sagittal position	Frontal position x Features extraction method	Sagittal position x Features extraction method
1	0,339	0,938	0,288	0,354	0,945	0,618
2	0,280	0,394	0,034	0,340	0,234	0,782
3	0,479	0,588	1,20E-15	0,843	0,584	0,894
4	0,003	0,091	8,29E-10	0,387	0,029	0,646
5	0,645	0,584	1,56E-11	0,600	0,829	0,445
6	0,345	0,679	5,36E-10	0,812	0,055	0,228
7	0,554	0,925	2,44E-11	0,623	0,428	0,404
8	0,249	0,570	5,75E-12	0,046	0,393	0,112
9	0,481	0,821	4,18E-13	0,032	0,153	0,318
10	0,902	0,701	5,81E-06	0,275	0,297	0,396
11	0,015	0,116	6,51E-17	0,486	0,136	0,096
12	0,821	0,283	8,72E-13	0,906	0,879	0,676
13	0,695	0,711	2,51E-13	0,809	0,223	0,393
14	0,623	0,575	3,50E-05	0,455	0,958	0,907
15	0,677	0,510	5,44E-14	0,347	0,450	0,800
16	0,399	0,159	1,05E-09	0,860	0,382	0,681
17	0,061	0,240	4,30E-08	0,489	0,052	0,814
18	0,782	0,036	4,94E-14	0,896	0,123	0,669
19	0,972	0,531	3,08E-04	0,348	0,677	0,624
20	0,513	0,204	5,33E-04	0,805	0,792	0,135
TOTAL	2	1	19	2	1	0

Tableau 5.4 – Linear kernel: a 3-factor analysis of variance (ANOVA) was performed on each subject for each ROI: Frontal position (3 levels: front, median and back), Sagittal position (3 levels: left, central and right), features extraction method (6 levels: roiGFP, roiMean, roi4bFFT, roiMean4bFFT, roiWhFFT and roiMeanWhFFT). The post-hoc comparison of inter-category interactions was performed with a Bonferroni correction. The significance of the results was set at $p < 0.05$.

Table 5.5 shows the statistical results of the 3-factor ANOVA performed for the RBF kernel. All of 20 subjects reach the significance threshold ($p < 0.001$) for the “features extraction method” factor. The pValue is significant for 1 subject for the “frontal position” factor and for 2 subjects for the “sagittal position” factor. The interactions between sagittal and frontal position, between frontal position and type of features, and between sagittal position and type of features were significant only for 1, 2 and 1 subjects respectively ($p < 0.05$). Again, when Bonferroni-correcting these analyses, we found that only the “features extraction method” factor reached the significance threshold ($\alpha < 0.025$).

Subject	Frontal position	Sagittal position	Features extraction method	Frontal position x Sagittal position	Frontal position x Features extraction method	Sagittal position x Features extraction method
1	0,744	0,214	3,08E-24	0,956	0,398	0,765
2	0,134	0,768	2,79E-21	0,305	0,357	0,605
3	0,648	0,919	1,75E-30	0,298	0,817	0,191
4	0,510	0,510	5,83E-23	0,222	0,632	0,230
5	0,899	0,011	4,23E-27	0,127	0,029	0,092
6	0,065	0,054	3,81E-18	0,238	0,123	0,623
7	0,007	0,257	1,69E-24	0,275	0,333	0,202
8	0,132	0,028	2,90E-27	0,002	0,005	0,029
9	0,450	0,443	1,11E-25	0,407	0,910	0,941
10	0,641	0,146	7,62E-20	0,518	0,076	0,108
11	0,368	0,457	1,35E-27	0,289	0,132	0,630
12	0,726	0,600	3,51E-23	0,850	0,842	0,924
13	0,789	0,765	6,96E-23	0,648	0,924	0,716
14	0,568	0,570	1,12E-18	0,597	0,313	0,493
15	0,749	0,540	9,62E-23	0,431	0,916	0,941
16	0,695	0,758	2,27E-20	0,422	0,516	0,885
17	0,991	0,576	8,09E-15	0,943	0,951	0,474
18	0,589	0,129	1,85E-26	0,121	0,720	0,185
19	0,789	0,834	4,84E-17	0,454	0,654	0,614
20	0,136	0,708	6,24E-05	0,820	0,882	0,843
TOTAL	1	2	20	1	2	1

Tableau 5.5 – RBF kernel: a 3-factor analysis of variance (ANOVA) was performed on each subject for each ROI: Frontal position (3 levels: front, median and back), Sagittal position (3 levels: left, central and right), features extraction method (6 levels: roiGFP, roiMean, roi4bFFT, roiMean4bFFT, roiWhFFT and roiMeanWhFFT). The post-hoc comparison of inter-category interactions was performed with a Bonferroni correction. The significance of the results was set at $p < 0.05$.

5.2.4.5/ EVOLUTION OF HYPERPARAMETERS

We explored the influence of the variation of the hyperparameters on the accuracy obtained, both on the training data and on the validation and test data.

Linear kernel

Figure 5.4 represents the evolution of the training, validation and testing accuracies as a function of the hyperparameter C for the linear kernel, according to the type of features used on all 257 electrodes. Figure 5.4 a represents the typical evolution for the use of the GFP, Figure 5.4 b for the use of the Fourier transform over 4 frequency bands, while Figure 5.4 c shows the variations in accuracy induced by the change in C for a complete Fourier transform. On the 3 graphs, the training accuracy is always higher than the validation accuracy, which itself is higher than the testing accuracy. Figure 5.4 a and 5.4 c show that as C decreases, training accuracy decreases and validation and testing accuracy increases, until they are almost equivalent around $C = 10^{-6}$ (depending on the subject ob-

served). These results illustrate the fact that the higher the C , the more rigid and complex the data separation boundary is during training and the less the model thus created has the ability to generalize on unknown data and to perform well on test data.

On Figure 5.4 b, on the contrary, the 3 accuracies stay equivalent between training, validation and testing, and that, moreover, these results remain remarkably stable in the face of C variations.

It is remarkable that all subjects, independently of the method or ROI used, systematically belong to one of these two profiles. Therefore, we have tallied the number of participants in each profile for each method considered (Figure 5.4 c). We thus show that only the extraction methods based on 4 frequency bands fall into profile B, and this for all subjects except 6 and 17 when using the 257 electrodes. All subjects for all other methods belong to profile A.

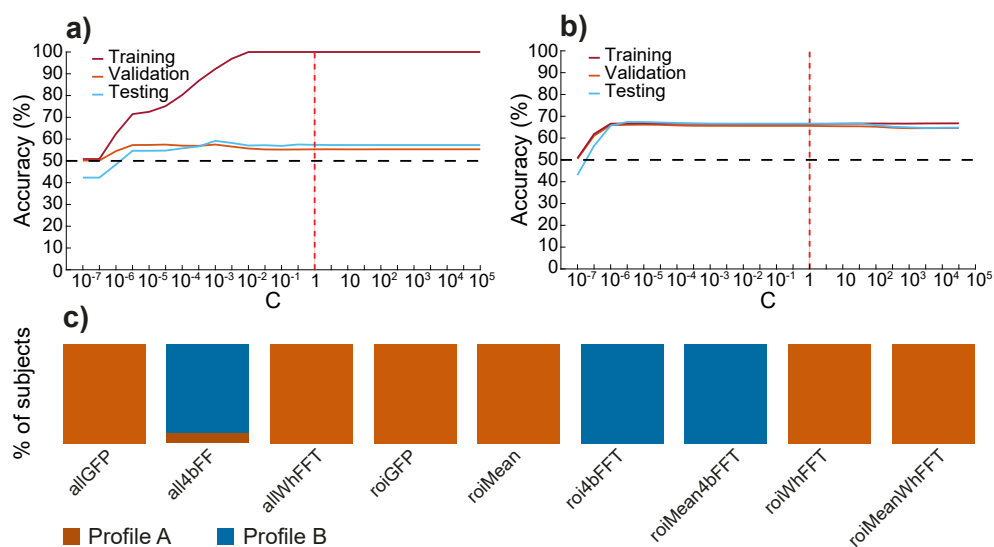


FIGURE 5.4 – Linear kernel: evolution of training, validation and testing accuracies as a function of the hyperparameter C . a) Profile A. b) Profile B. c) Number of subjects who belong to profile A or B for each features extraction method.

RBF kernel

Two different profiles can be distinguished by looking at the heatmaps obtained after gridsearch of the hyperparameters (Figure 5.5).

The accuracy of training (Figure 5.5 a and b), validation (Figure 5.5 c and d) and testing (Figure 5.5 e and f) is plotted as a function of C and sigma for the RBF kernel. Accuracy is represented by the colorbar, ranging from 50 to 100% accuracy.

Heatmaps in Figures 5.5 c and e, and in Figure 5.5 d and f are almost identical, meaning that the validation and test accuracies are very similar for each profile. However, heatmaps in Figures 5.5 a and b show that the accuracy obtained during training is generally higher, and even punctually reaching 100% for some pairs of hyperparameters.

On profile A, as much as the testing accuracy (Figure 5.5 e) varies significantly as a function of C , for a given C , few variations as a function of sigma can be observed. Additionally, the standard pair ($C = 1$, sigma = 1) gives a result close to random, but by keeping

the standard C and varying the sigma, better results are obtained, with a great stability between $\sigma = 50$ and $\sigma = 1000$ (0.05% difference only).

On profile B, the variation of the hyperparameters does not give a significant change in validation and testing accuracy (Figures 5.5 d and f), which remain close to 50% over the entire grid.

All subjects, regardless of the method and ROI considered, systematically fall into one or the other of these profiles. We have therefore counted the number of subjects that correspond to either profile (Figure 5.5 g). Concerning the features based on the time domain (GFP all elec, ROI mean and ROI GFP), 14 subjects correspond to profile B, displaying a test accuracy close to 50%, which does not increase with the hyperparameters. For the methods using the Fourier transform on 4 frequency bands, 15 subjects belong here to profile A, with an improvement of the test accuracy on part of the hyperparameter pairs. Finally, for methods using the whole Fourier transform, the subjects are spread equally on profile A and B.

5.2.4.6/ DENSITY OF HYPERPARAMETERS

Linear kernel

We examined the results achieved based on the best validation accuracy for each of the 50 permutations. However, the hyperparameters selected as the best are not necessarily the same each time. We wondered how often each pair of hyperparameters was selected, and whether one or more of them came out more frequently. Results for linear and RBF kernel, and for the following methods are given in Figures 5.6, 5.7 and 5.8 and in Figures 5.9, 5.10 and 5.11 respectively: roiMean, roi4bFFT and all4bFFT.

The sum of occurrences for which a hyperparameter was chosen over all 50 permutations for all subjects is shown as a histogram for the linear kernel. In Figure 5.6 a and the roiMean method, we show that the preferred C values are very low, between 10^{-5} and 10^{-6} , and that for most of the subjects the dispersion is low (Figure 5.6 b). The standard C = 1 is hardly chosen, and hyperparameters higher than 10 are non-existent.

In Figure 5.7 a, for the roi4bFFT method, two peaks can be observed, one again around 10^{-6} , and another wider one between 10^{-3} and 1. Values of C above 10 are again not chosen. Interestingly, the subject-by-subject dispersion shows two groups of subjects with a small standard deviation for each of these peaks, and 5 subjects with a large dispersion from one peak to its neighbour (Figure 5.7 b).

In Figure 5.8 a, on the other hand, the same two peaks as in Figure 5.7 a can be observed, but while C values close to 1 are very sparsely chosen, this is not the case for C greater than 10, where an increase in the sum across subjects is observed from $C = 10^{-2}$ onwards. This phenomenon is due to subjects 5, 7, 10, 14 and 18, whose dispersion is extremely high, as well as to subjects 6 and 17, whose standard deviation is low but whose mean is concentrated on very high C values (Figure 5.8 b).

RBF kernel

We performed the same analysis on the same feature extraction methods on the hyperparameters of the RBF kernel (Figures 5.9, 5.10 and 5.11).

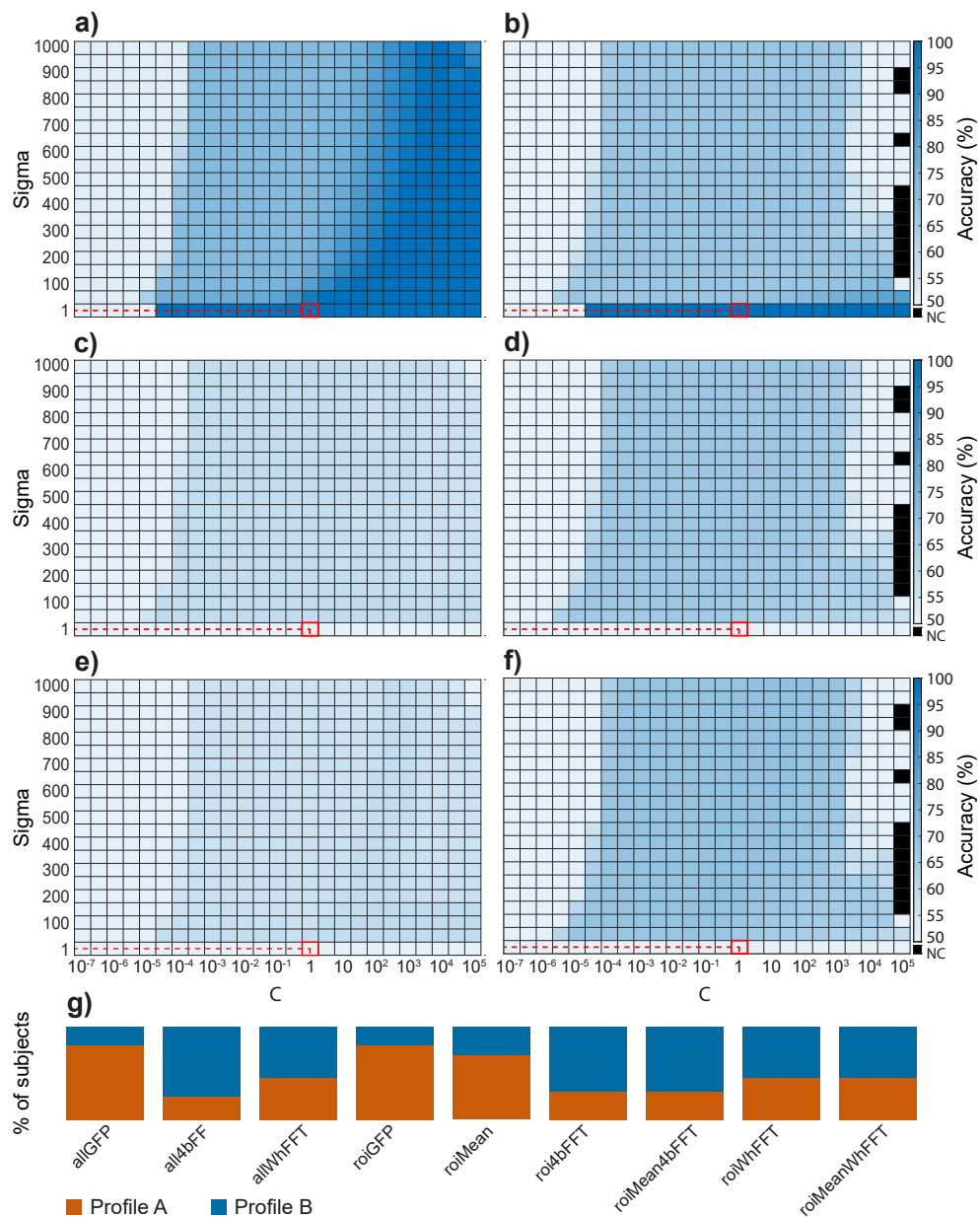


FIGURE 5.5 – RBF kernel: evolution of training, validation and testing accuracies as a function of the hyperparameters C and sigma. a) Training accuracy, profile A. b) Training accuracy, profile B. c) Validation accuracy, profile A. d) Validation accuracy, profile B. e) Testing accuracy, profile A. f) Testing accuracy, profile B. g) Number of subjects who belong to profile A or B for each features extraction method.

A large number of combinations is never chosen. For C between 10^{-4} and 10^{-3} , any sigma can be chosen to form the pair. This is also the case in Figures 5.10 b and 5.11 b for C between 10^2 and 10^3 . For Figure 5.9 b, on the other hand, this phenomenon is much more spread out and runs from C = 1 to C = 10^5 . This spiking density phenomenon on some C is similar to the one we have seen for C in the linear kernel.

It is worth noting that the most selected combinations always have a sigma less than or equal to 50, and are distributed over the two peaks of C.

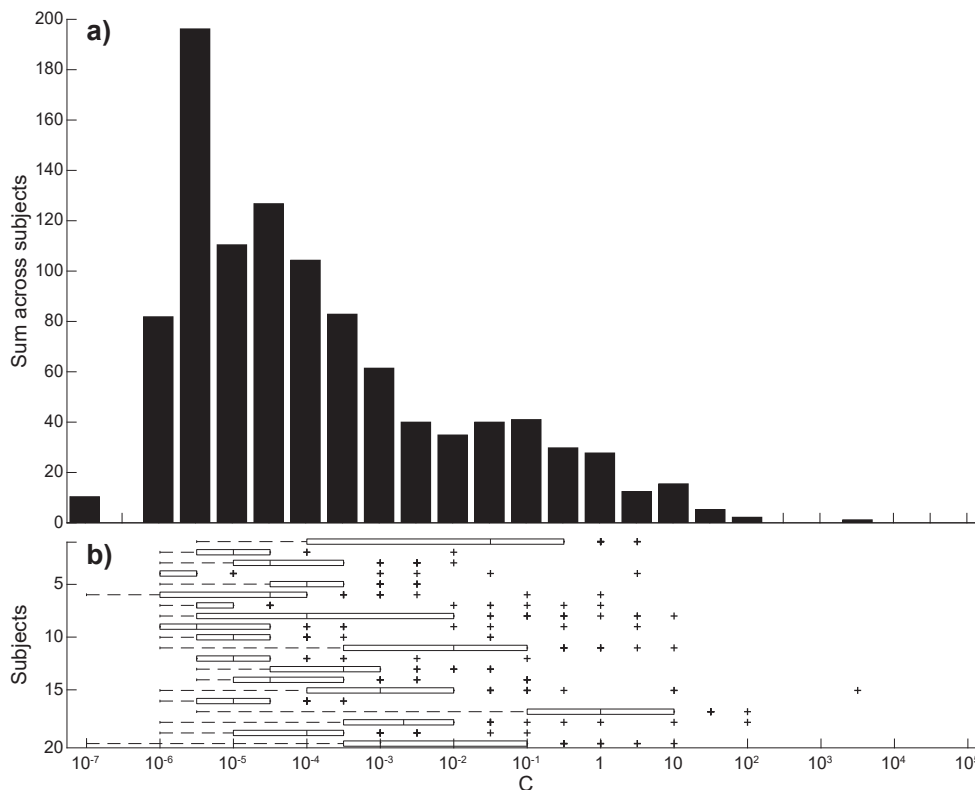


FIGURE 5.6 – Linear kernel: density of chosen hyperparameter C across 50 permutations, for roiMean method. a) Number of times a hyperparameter was chosen over all 50 permutations for all subjects. b) Number of times a hyperparameter was chosen over all 50 permutations for each subject individually.

Figures 5.9 a, 5.10 a and 5.11 a plot the average sigma density subject by subject, and show that for most subjects, the dispersion is low and is rather concentrated on sigmas below 200.

Conversely, Figures 5.9 c, 5.10 c and 5.11 c describe subjects with a high dispersion of C density between 10^{-4} and 10^4 .

5.2.4.7/ COMPARISON WITH STANDARD HYPERPARAMETERS

Since we are looking for the best hyperparameters for our classification problem, it makes sense to compare the results then acquired with those obtained using standard hyperparameters, *i.e.* $C = 1$ for the linear kernel, and $C = 1$ and $\sigma = 1$ for the RBF kernel (Table 5.6, Table 5.7).

On average, the difference is 0.5% (SD = 1.4%; range = [-1.2% - 3.7%]) for the linear kernel, and 4.8% (SD = 3.5%; range = [-0.8% - 13.7%]) for the RBF kernel. In general, standard hyperparameters produce a lower accuracy than optimised hyperparameters, and the difference is greater for the RBF kernel than for the linear. This tends to corroborate the need for tuning kernel parameters. However, for some subjects, the accuracy is higher with standard hyperparameters. Although this may seem counter-intuitive, it is understandable given that the best hyperparameters are selected based on validation ac-

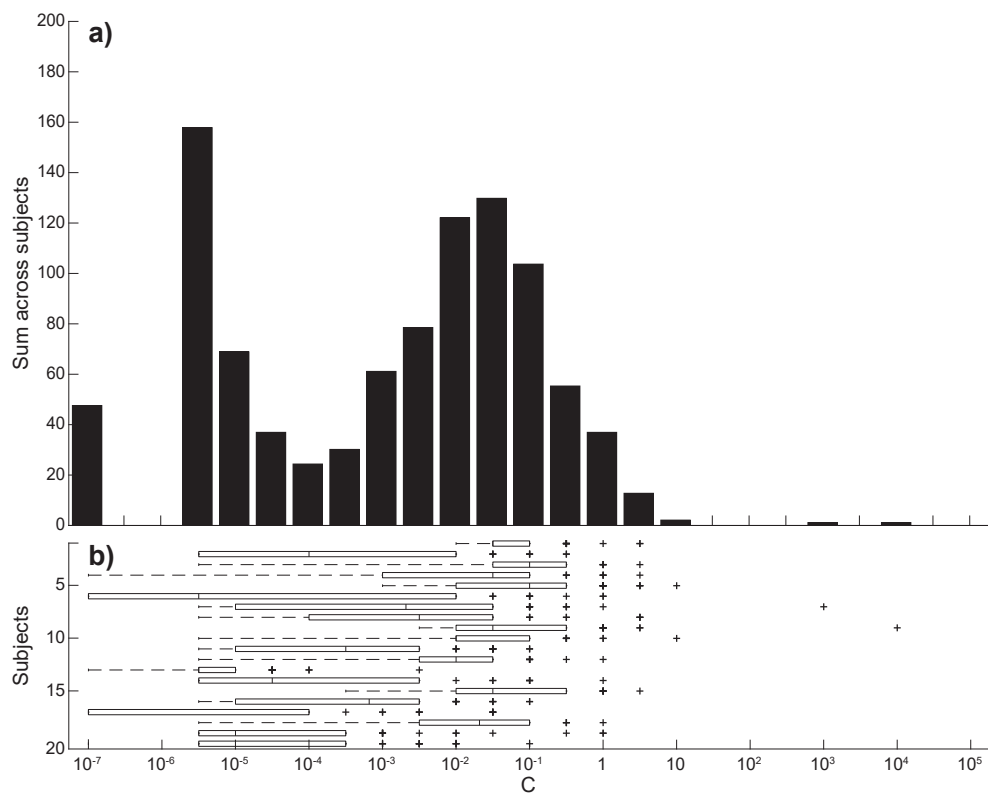


FIGURE 5.7 – Linear kernel: density of chosen hyperparameter C across 50 permutations, for roi4bFFT method. a) Number of times a hyperparameter was chosen over all 50 permutations for all subjects. b) Number of times a hyperparameter was chosen over all 50 permutations for each subject individually.

curacy. Indeed, the best hyperparameter pair based on the performance of the validation set does not guarantee good performance on the test data, highlighting heterogeneity in our dataset.

5.2.4.8/ COMPARISON WITH BEST HYPERPARAMETERS BASED ON TEST

It is easy, when using cross-validation in classification for features or hyperparameters optimization, to make the mistake of basing our selection on the test result. We chose to examine the results obtained by this means and compare them, kernel by kernel, with those obtained by the rigorous method we presented above (Table 5.6, Table 5.7).

For the linear kernel, mean performance was $X = 79.5\%$ of accuracy (SD = 5.2%; range = [68.3% - 89.7%]) (Table 5.6), and for the RBF kernel it was $X = 79.3\%$ (SD = 9.2%; range = [54.4% - 96.0%]) (Table 5.7). We observe that the accuracy based on the test is for all subjects (except 19 and 20 with the RBF) much higher than the accuracy based on the validation: there is a difference of 17.7% on average for linear kernel (SD = 3.2%; range = [9.4% - 21.7%]), and of 16.7% on average for RBF kernel (SD = 7.0%; range = [-2.0% - 25.9%]).

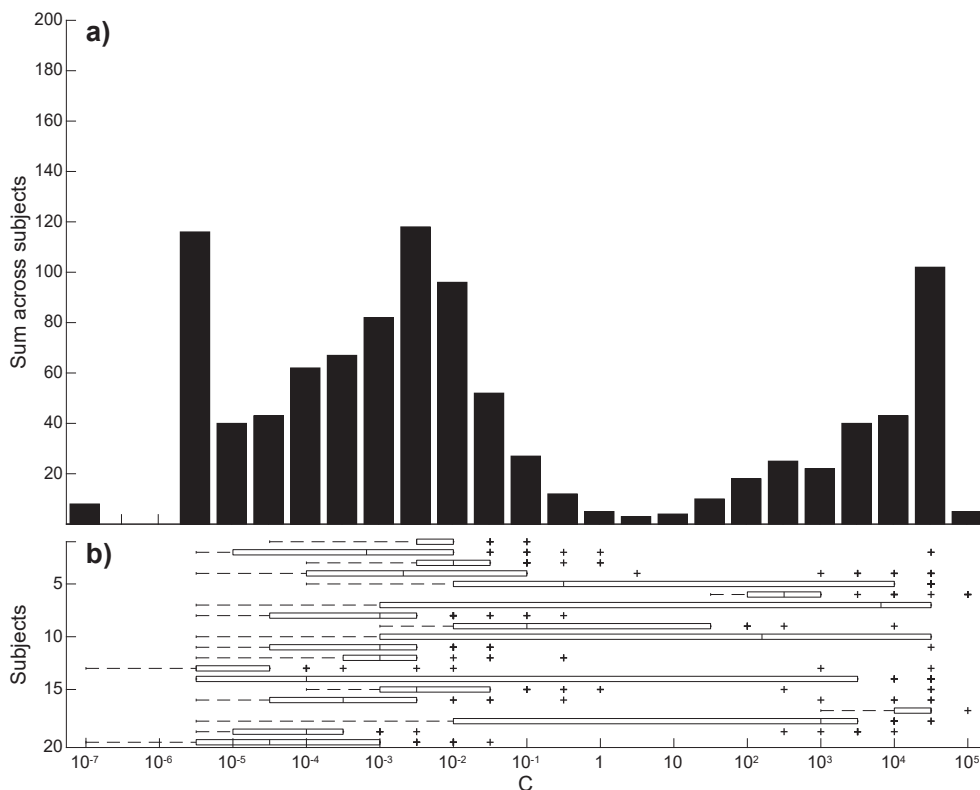


FIGURE 5.8 – Linear kernel: density of chosen hyperparameter C across 50 permutations, for all4bFFT method. a) Number of times a C value was chosen over all 50 permutations for all subjects. b) Number of times a C value was chosen over all 50 permutations for each subject individually.

5.2.4.9/ VERIFICATION OF RESULTS WITH RANDOMLY GENERATED DATA

In order to verify that our classification pipeline does not obtain these results on our data by chance, we have reproduced our classification scheme on randomly generated data. We therefore generated a dataset for each feature extraction method used, in such a way as to have the same number of features each time. The results are summarized in Table 5.8, and show that the performance of the random data is much lower than that of our data and does not exceed 50% accuracy, both with optimized and standard hyperparameters.

However, when looking at the accuracy based on the test, the random data are able to achieve up to 62.2% accuracy (RBF kernel, roiMean4bFFT), demonstrating that method error in cross-validation is to be avoided at all costs.

Interestingly, the feature extraction method that obtains the best performance regardless of the kernel or hyperparameters considered is roiMean4bFFT, *i.e.* the one with the shortest feature vectors - only 4 features.

5.2.4.10/ RESULTS OF PERMUTATION TESTS

In parallel with the Müller-Putz's evaluation of the performance reliability [353], we conducted statistical tests of labels permutation. If the accuracy obtained with random class labels is higher than that obtained with real labels, then the performance of the

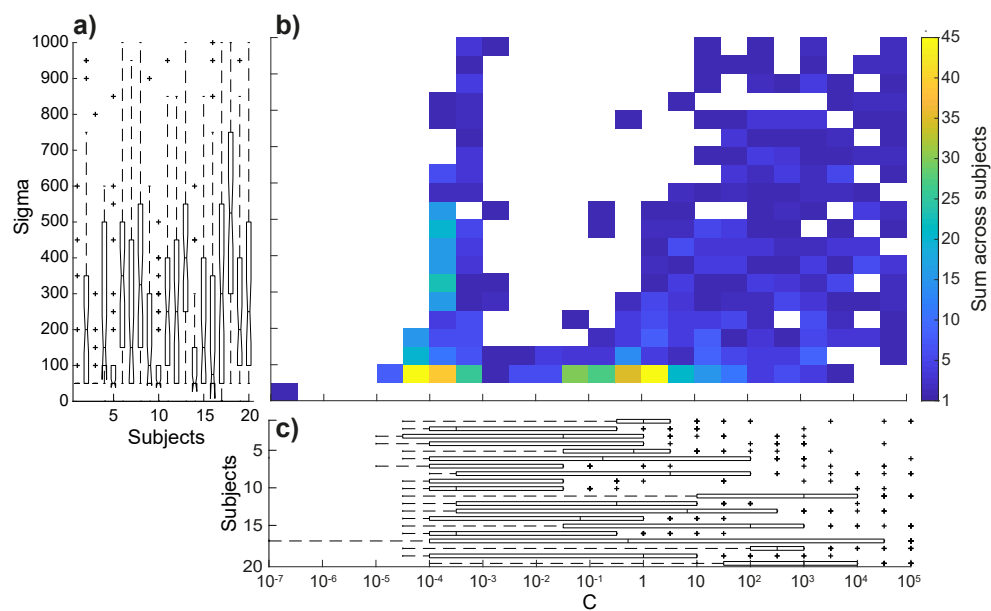


FIGURE 5.9 – RBF kernel: density of chosen pair of hyperparameters C and sigma across 50 permutations, for roiMean method. a) Number of times a sigma value was chosen over all 50 permutations for each subject individually. b) Number of times a pair (C, sigma) was chosen over all 50 permutations for all subjects. c) Number of times a C value was chosen over all 50 permutations for each subject individually.

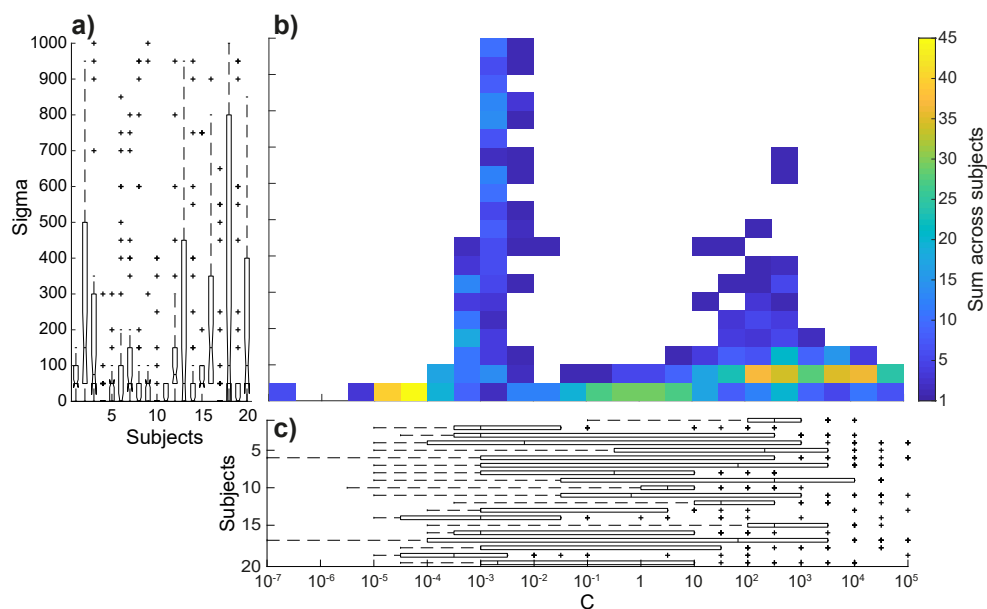


FIGURE 5.10 – RBF kernel: density of chosen hyperparameter C across 50 permutations, for roi4bFFT method. a) Number of times a sigma value was chosen over all 50 permutations for each subject individually. b) Number of times a pair (C, sigma) was chosen over all 50 permutations for all subjects. c) Number of times a C value was chosen over all 50 permutations for each subject individually.

classifier is not reliable. These results are given as pValue in Tables 5.2 and 5.8. They are similar to those given by the Müller-Putz method, except for subjects 17 and 20 for

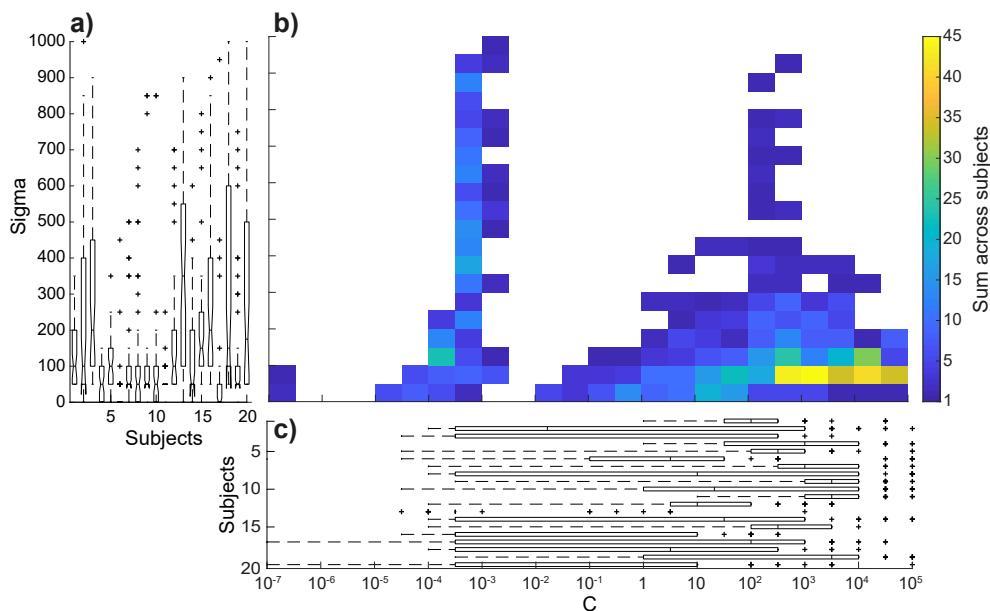


FIGURE 5.11 – RBF kernel: density of chosen hyperparameter C across 50 permutations, for all4bFFT method. a) Number of times a sigma value was chosen over all 50 permutations for each subject individually. b) Number of times a pair (C, sigma) was chosen over all 50 permutations for all subjects. c) Number of times a C value was chosen over all 50 permutations for each subject individually.

the linear kernel, and subjects 6 and 17 for the RBF kernel.

5.2.5/ DISCUSSION

This study aimed to investigate the possibility of classifying EEG data from musical imagery. We explored 9 features extraction methods and the influence of electrode selection first, and then the influence of the SVM architecture employed on performance, subject by subject.

5.2.5.1/ INFLUENCE OF THE SELECTED ROI

Thanks to the results of the 3-factor ANOVA, we showed, for both the linear and RBF kernels, what was the respective influence of the position of the ROI on the scalp and of the feature extraction method on the accuracies obtained. For the two kernels, all subjects present a highly significant pValue for the “features extraction method” factor. On the other hand, it appears that the position of the ROI, whether frontal or sagittal, has no significant influence, as after a post-hoc Bonferroni correction, no pValue is significant anymore.

This observation goes against our hypothesis that some ROIs would perform better than others. Indeed, in Gabriel *et al.* (2016) [292], the temporal and central ROIs showed a significant difference in amplitude, correlated with the sources of brain activity (temporal areas and SMA). While we expected that at the individual level the ROIs concerned might differ from the group average, and from one subject to another, it is surprising that no ROI emerges from the whole. Indeed, it is highly unlikely that all brain areas are activated in

Subject	<u>Linear kernel</u>					
	Based on validation	Standard hyperparameters			Based on test	
	Accuracy	Accuracy	Difference val-standard	Accuracy	Difference test-val	
1	70,6%	69,4%	1,2%	80,0%	9,4%	
2	59,8%	56,2%	3,6%	79,3%	19,5%	
3	69,4%	69,5%	-0,1%	88,0%	18,6%	
4	63,3%	59,7%	3,7%	79,3%	16,0%	
5	64,0%	65,0%	-1,0%	83,3%	19,3%	
6	54,6%	53,3%	1,3%	68,3%	13,7%	
7	61,1%	61,5%	-0,4%	82,8%	21,7%	
8	59,3%	60,6%	-1,2%	75,9%	16,6%	
9	68,1%	68,6%	-0,4%	89,7%	21,5%	
10	65,0%	65,1%	-0,1%	81,0%	16,0%	
11	58,7%	58,2%	0,5%	78,8%	20,1%	
12	64,2%	63,2%	1,0%	84,8%	20,7%	
13	57,2%	55,7%	1,5%	75,8%	18,6%	
14	58,5%	56,8%	1,7%	78,8%	20,3%	
15	58,6%	59,2%	-0,7%	76,2%	17,6%	
16	69,0%	67,4%	1,6%	81,0%	12,0%	
17	55,3%	56,2%	-1,0%	71,4%	16,1%	
18	61,4%	62,0%	-0,5%	81,0%	19,5%	
19	61,2%	60,4%	0,7%	81,8%	20,7%	
20	55,9%	56,5%	-0,6%	72,7%	16,8%	
Mean	61,8%	61,2%	0,5%	79,5%	17,7%	
SD	4,8%	4,9%	1,4%	5,2%	3,2%	
Min	54,6%	53,3%	-1,2%	68,3%	9,4%	
Max	70,6%	69,5%	3,7%	89,7%	21,7%	

Tableau 5.6 – Linear kernel, for each subject: best testing accuracy based on validation, testing accuracy based on standard $C = 1$ and best testing accuracy based on test.

the same way between the two conditions for all subjects, when the group study shows differently.

On the other hand, the fact that the feature extraction method has an influence and that this influence is identical for both kernels indicates a certain robustness of the data and features from one kernel to the other.

5.2.5.2/ IMPACT OF THE FEATURES EXTRACTION METHOD ON PERFORMANCE

Although the position of the ROI does not seem to be critical here, the use of one ROI or all 257 electrodes does affect the performance. Indeed, methods using ROI are more prevalent when looking for the best performing method. Having fewer features would be more efficient, and would therefore favour in theory ROI 4b mean methods which only have 4. This hypothesis is logical if we consider that a larger number of trials than features is required to create an optimal decision frontier between the two classes [354, 355].

But music is a complex stimulus, involving high level cognitive processes [356]. EEG data obtained by this process are therefore complex and shaded, and with small number of

Subject	RBF kernel					
	Based on validation	Standard hyperparameters			Based on test	
	Accuracy	Accuracy	Difference val-standard	Accuracy	Difference test-val	
1	68,0%	65,0%	3,0%	92,0%	24,0%	
2	60,8%	51,0%	9,8%	82,8%	21,9%	
3	70,1%	63,2%	6,9%	96,0%	25,9%	
4	62,4%	55,7%	6,7%	79,3%	16,9%	
5	64,2%	62,3%	1,9%	85,7%	21,6%	
6	55,1%	52,7%	2,4%	70,7%	15,6%	
7	61,9%	58,1%	3,8%	82,8%	20,9%	
8	62,6%	61,2%	1,4%	79,3%	16,7%	
9	69,1%	66,5%	2,7%	82,8%	13,6%	
10	65,7%	66,5%	-0,8%	85,7%	20,0%	
11	62,3%	61,2%	1,1%	78,8%	16,5%	
12	63,9%	57,5%	6,4%	81,8%	17,9%	
13	60,3%	51,0%	9,3%	78,8%	18,4%	
14	59,7%	55,9%	3,7%	78,8%	19,1%	
15	59,6%	54,7%	4,8%	81,0%	21,4%	
16	68,9%	62,3%	6,6%	83,3%	14,4%	
17	54,6%	50,0%	4,6%	73,8%	19,2%	
18	62,6%	57,3%	5,2%	76,2%	13,6%	
19	63,4%	49,7%	13,7%	62,5%	-0,9%	
20	56,5%	54,4%	2,0%	54,4%	-2,0%	
Mean	62,6%	57,8%	4,8%	79,3%	16,7%	
SD	4,4%	5,5%	3,5%	9,2%	7,0%	
Min	54,6%	49,7%	-0,8%	54,4%	-2,0%	
Max	70,1%	66,5%	13,7%	96,0%	25,9%	

Tableau 5.7 – RBF kernel, for each subject: best testing accuracy based on validation, testing accuracy based on standard $C = 1$ and $\sigma = 1$, and best testing accuracy based on test.

trials it is difficult to represent them. It is the 4-band features on ROI (mean of electrodes and electrode by electrode), together with the average ROI that are the best method for the majority (Figure 3), *i.e.* 2/3 of the subjects. This partly contradicts our hypothesis that it is better to work with a reduced number of features, because we move from 4 and 32 features to 1900 respectively, and methods with a number of features between these limits represent a minority. So the performance is not directly proportional to the number of features.

This shows that while the frequency content is more relevant for some subjects, for others the essential information is better embedded in the temporal signal [357].

As there is a high degree of inter-individual variability, we expected that the optimal methods would change from one individual to another. However, sometimes the difference in performance between the best method and the second one for the same subject is not large, and here the size of this difference is not taken into account.

Kernel	Features	Based on validation	Standard hyperparameters	Based on test
Linear	allGFP	42,9%	49,2%	53,3%
	all4bFFT	43,0%	49,5%	53,0%
	allWhFFT	41,7%	47,1%	50,6%
	roiGFP	42,9%	49,2%	53,3%
	roiMean	42,9%	49,2%	53,3%
	roi4bFFT	42,1%	44,4%	51,2%
	roiMean4bFFT	50,2%	54,2%	56,6%
	roiWhFFT	43,0%	46,6%	53,7%
	roiMeanWhFFT	49,6%	52,2%	59,2%
	Best method	roiMean4bFFT	roiMean4bFFT	roiMeanWhFFT
RBF	allGFP	51,9%	42,0%	55,1%
	all4bFFT	52,7%	42,0%	55,6%
	allWhFFT	51,3%	44,6%	52,4%
	roiGFP	53,9%	42,2%	57,7%
	roiMean	51,9%	42,0%	55,1%
	roi4bFFT	51,9%	42,0%	55,1%
	roiMean4bFFT	55,4%	51,5%	62,2%
	roiWhFFT	53,0%	41,6%	56,0%
	roiMeanWhFFT	54,7%	42,6%	60,5%
	Best method	roiMean4bFFT	roiMean4bFFT	roiMean4bFFT

Tableau 5.8 – Results of classification pipeline on randomly generated data, for both linear and RBF kernels, for each features extraction method. Size of features vectors were preserved for each features extraction method for fair comparison. Best testing accuracy based on validation, testing accuracy based on standard hyperparameters and best testing accuracy based on test.

5.2.5.3/ SVM ARCHITECTURE: INFLUENCE OF KERNELS, HYPERPARAMETERS AND CROSSVALIDATION

The methods using 4 frequency bands present a plateau in the evolution of the hyperparameters according to profile B, for both linear and RBF kernels. This reveals a high stability of the model with respect to hyperparameter variations.

Furthermore, it shows that the selected hyperparameters are reliable, as the presence of the plateau guarantees the absence of any accuracy peak potentially present between the two combinations examined.

Thus, the 4-band methods not only provide better classification accuracy, but are also more robust. They could represent a more reliable way to carry out classification by ensuring that the correct accuracy is not accidental. The same observation can be made with the RBF kernel.

We observe a variation in the hyperparameter density for the top 3 feature extraction methods. Overall the most selected hyperparameters are chosen in the smaller C (10^{-6} to 10^{-2}). The standard C is rarely chosen. Despite the robustness of the methods examined to hyperparameters variations, picking the best hyperparameter at each permutation may provide a better amortization of the data heterogeneity problem. This information supports

the need for hyperparameters to be optimised.

Regarding the RBF kernel, operating on two hyperparameters instead of just one may result in different combinations leading to the same SVM model. The peak effect observed on the 3 feature extraction methods suggests that C influences performance more than σ , partly matching the phenomenon observed for the linear kernel. A standard σ may be sufficient on our data, and that may allow us to tune the C only, which would save computing time. Another possibility would be to explore a smaller σ range, for example between 1 and 50.

5.2.5.4/ STANDARD HYPERPARAMETERS

If standard hyperparameters are selected for each permutation (without optimisation), the difference with the result optimised for the linear kernel is very slight, and sometimes even the standard parameters are more efficient. A difference of 5% on average and 14% maximum is observed for the RBF kernel. Therefore, optimising the hyperparameters appears to be more essential for the RBF kernel than for the linear kernel.

Theoretically, RBF should give better results, for our data are complex and difficult to linearly separate: 2 hyperparameters are available to tune instead of one, making the method more flexible. Best test The performance improvement is extremely substantial when using this cross-validation method. But although very satisfying in terms of accuracy, this approach is a methodological error. We would like to emphasize that this error provides fairly credible numbers, and is therefore difficult for the novice to recognise.

5.2.5.5/ RANDOMLY GENERATED DATA

The results obtained with random data do not exceed 55.4% accuracy for both the linear and RBF kernel for the optimization best accuracy based on validation and standard hyperparameters.

In contrast, with the cross-validation best test, we reach 65%, so RBF can classify between noise from random data. We thus recommend a systematic comparison between the results achieved on the dataset and those obtained on random data to ensure that statistical noise is indeed eliminated.

5.2.5.6/ STATISTICAL ANALYSIS OF CLASSIFICATION ACCURACIES

We applied 2 different approaches: test permutation and Müller-Putz's threshold [353]. Subjects 6, 17 and 20 are not statistically significant, but we expected them to be negative because their knowledge of their familiar songs was insufficient, thus reducing the difference between well-known and new songs. The two methods always agree, except for these 3 subjects. This demonstrates first of all that the SVM, even after optimisation, is not able to classify subjects for whom the protocol did not really succeed, *i.e.* it does not make a classification on the background noise. This confirms that our pipeline is reliable. Moreover, the discrepancy between the statistical performance evaluation methods for these subjects shows that these data are not purely random, but rather borderline cases.

17 out of 20 subjects significantly emerge from statistical noise. In any case, despite the

optimisation of the hyperparameters, the pValue is never significant for random data, once again confirming the pipeline we propose is robust against overfitting.

As a general matter, our data are quite heterogeneous, with high intra-class variability. The reason for this is that we chose a protocol that reflects what happens in natural conditions when people are confronted with an unknown song, compared to songs that are very familiar to them. The songs are therefore all different, and variations in the stimuli are consequently present (tempo, human voice, emotion conveyed, *etc...*). Moreover, within the same song, each stimulus is unique because the silent periods are scattered at different moments in the song. The complexity of our data largely explains why the accuracy remains relatively low compared to what can be observed in the literature: the training data are not properly representative of the test data, so the result obtained on these two sets is uneven.

Finally, the RBF kernel is a little more performant in terms of accuracy than the linear one, but not more efficient because it is very long. Under these circumstances, it would be more relevant to consider the linear kernel. Concomitant work from Wainer and Fonseca (2020) [358] comparing 18 different methods of hyperparameters tuning for RBF is also an interesting lead to increase performance of RBF kernel in future works.

5.3/ CONCLUSION

In this study, we aimed at classify EEG recordings of brain activity during musical imagery induced by silences inserted in songs. We used features from the Fourier Transform or from 4 frequency bands of interest, and from Global Field Power to compare their relevance subject by subject. We also considered the use of all electrodes or individual regions of interest. Classification was performed by SVM, and we tuned the hyperparameters of two different kernels by nested crossvalidation and gridsearch.

Results showed that while the use of ROI give best results than all electrodes for most subjects, the spatial position of those is not important for classification performance. Features extraction methods involving the 4 frequency bands reached higher accuracies and were more stable against hyperparameters variations for both linear and RBF kernels, thus more reliable for classification. While both kernels provided comparable accuracies, a way shorter computation time for linear kernel makes it more attractive.

Finally, comparison with randomly generated data and statistical tests results demonstrate that our pipeline is robust and reliable for this kind of data.



ÉTATS DE CONSCIENCE PATHOLOGIQUES : LE PROJET COMÉTIQUE

LE TRAITEMENT SÉMANTIQUE COMME INDICATEUR DE CONSCIENCE RÉSIDUELLE

6.1/ INTRODUCTION

Les états de conscience altérée témoignent d'un dysfonctionnement cérébral sévère pouvant avoir diverses origines : le traumatisme crânien créant une hémorragie cérébrale est la plus courante, mais un trouble métabolique faisant suite à une hypoglycémie sévère ou à l'exposition à des agents toxiques, tels que drogues, médicaments, insecticides ou alcool peut également en être responsable, tout comme une infection ou une tumeur. Enfin, la cause du coma peut être d'origine hypoxique ou anoxique lorsque le cerveau a manqué d'oxygène, soit à cause d'un arrêt cardio-circulatoire (infarctus du myocarde, trouble du rythme cardiaque) soit à la suite d'une insuffisance respiratoire aiguë (bronchopneumopathie chronique et de surinfection ou pneumopathie d'inhalation).

Bien que l'EEG conventionnelle soit couramment employée dans la pratique clinique, le faible nombre d'électrodes utilisées pose un problème majeur qui est celui de leur placement : l'ERP recherché peut être présent mais non détecté à cause d'un positionnement des électrodes mal adapté au sujet. Il est donc impossible de conclure à l'absence réelle de l'ERP. Or, une forte corrélation a pu être établie entre la présence de certains ERPs cognitifs et la récupération du patient [359]. Ainsi, la détection d'ERPs tels que la N100, la MMN, la P300 et la N400 est annonciatrice de meilleures chances de rétablissement : 75% des patients qui décèdent ne présentaient pas ces ondes [360]. Toutefois, elles n'ont pu être enregistrées que sur 58,5% des patients ayant récupéré seulement, traduisant un taux élevé de faux-négatifs : l'absence de détection ne signifiant pas nécessairement la détection formelle de l'absence d'ERP.

L'étude Cométique est un essai clinique dont l'objectif principal est d'étudier si l'EEG-HR augmente la fiabilité de la prédiction de récupération des patients dans le coma par rapport à l'EEG normal. L'utilisation de l'EEG-HR se justifie donc par la possibilité d'obtenir un tableau clinique de l'activité électrique sur l'ensemble du scalp et non plus seulement sur des régions d'intérêt définies selon la littérature, grâce à un maillage d'électrodes serré et régulier [361–363]. Cet essai est concentré sur les patients en état de coma non traumatique issu d'une anoxie. En effet, l'EEG-HR comme outil de diagnostic est davantage adapté à ce type de coma, la pose d'électrodes pouvant poser des difficultés chez

certaines patients ayant des lésions au niveau du scalp suite à un traumatisme d'une part, et les structures cérébrales internes sont moins fréquemment endommagées, permettant la localisation des sources d'activité cérébrale d'autre part.

L'évaluation de la conscience dans cet essai est divisé en différentes sous tâches, pour une évaluation du niveau de conscience par approche hiérarchique :

- Étape 1 : évaluation de la capacité de traitement de l'information auditive et pré-attentionnelle, via l'élicitation des ERPs N100 et MMN. Le but est de déterminer en premier lieu si le sujet perçoit les sons, et est capable de les différencier. Il a donc été donné à écouter au sujet, de manière passive, deux types de stimuli auditifs, un son pur de fréquence 1000 Hz, fréquent, et un son pur à 2000 Hz, qualifié de "déviant" (15% des sons).

N100 : Sa présence chez les patients dans le coma suggère que les voies auditives sont fonctionnellement préservées. Certaines études indiquent que l'absence de l'onde N100 est un signe de mauvaise récupération [364, 365]. La présence d'une onde N100 est parfois considérée comme prédictif de récupération de la conscience [366, 367], mais pas toujours [365, 368].

MMN : Chez les patients dans le coma, la présence de la MMN est signe de récupération dans l'étiologie anoxique [369–372]. Bien que cet ERP soit plus souvent observé chez les patients UWS que chez les patients MCS (65% vs 34% des patients respectivement) [373–375] ont montré une augmentation de l'amplitude de la MMN lors de la transition UWS vers MCS.

- Étape 2 : évaluation de la capacité langagière et attentionnelle du sujet, via l'onde P300. Il a été présenté au sujet huit stimuli auditifs : le prénom du sujet, un autre prénom cible, et 6 autres prénoms non familiers, afin d'établir si l'attention du participant peut se porter automatiquement ou non vers un stimulus familier.

P300 : Chez les patients dans le coma, il est impossible de savoir si la capacité à suivre des instructions verbales est préservée, et donc si l'allocation des ressources attentionnelles est consciente ou non. La majorité de la littérature sur P300 et coma ne dissocie donc pas les composantes P3a et P3b et conserve la terminologie P300. Dans l'optique de pallier cette limitation, il était demandé à des patients en MCS et UWS de compter activement le nombre de fois qu'ils entendaient leur prénom ou un autre prénom [376]. Cette étude a montré que l'onde P300 des patients MCS était d'amplitude similaire à celle de sujets sains, alors qu'aucune onde n'était détectée chez les patients UWS.

- Étape 3 : évaluation de la capacité de traitement pré-sémantique, grâce à la N400. L'objectif est de caractériser la compréhension du sujet des sons perçus. 120 quatuors de mots ont ainsi été présentés, les trois premiers mots définissant un contexte sémantique, et le quatrième étant congruent ou non à ce contexte.

N400 : Bien que l'intérêt pour la N400 soit relativement nouveau dans le cadre de l'étude des états de conscience, des travaux menés chez des patients en UWS ou en MCS suggèrent que certains processus sémantiques pourraient être préservés [377]. En ce qui concerne le stade précoce de coma, Daltrozzo *et al.* (2009) [378] ont montré qu'un certain nombre de patients présentait une onde de type N400 lorsque des paires de mots sémantiquement incongruentes étaient présentées.

Notre objectif est d'évaluer la réponse sémantique obtenue lors d'enregistrements EEG sur sujets sains pour l'étape 3 de cette étude. Nous effectuons dans un premier temps une analyse de groupe conventionnelle, puis le second objectif est de classer l'activité cérébrale au niveau individuel en fonction de la congruence des sons perçus.

6.2/ MATÉRIEL ET MÉTHODES

6.2.1/ PROTOCOLE EXPÉRIMENTAL

Le paradigme utilisé pour induire la N400 repose sur l'utilisation de 120 quatuors de mots. Dans tous les cas, les trois premiers mots des quatuors se rapportent au même champ sémantique et servent à instaurer un contexte. Le quatrième et dernier mot est également congruent à ce contexte pour la moitié de ces quatuors, tandis qu'il est incongruent pour les soixante autres, ayant été permuté d'un quatuor à un autre (Figure 6.1). En effet, la présence d'un mot ne se rapportant pas au contexte mis en place crée une discordance cognitive observable sous la forme d'une onde N400 en EEG. Ce quatrième mot a été choisi de façon à ce que sa prononciation ne puisse pas induire une confusion quant à son sens (tel que "mal" et "mâle") et sa durée est systématiquement comprise entre 500 et 950 ms. L'espace entre deux mots d'un même quatuor était de 800 ms et deux quatuors différents étaient séparés par un intervalle de 1100 ms. Enfin, la présentation des quatuors a été randomisée sous la forme d'une liste unique, ayant donc un ordre de passage identique pour tous les sujets.

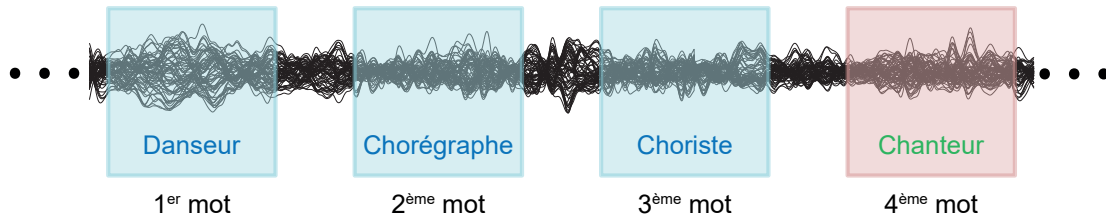
6.2.2/ PARTICIPANTS

Trente sujets droitiers en bonne santé ont été inclus dans cette étude et ont reçu une compensation financière. Les données de l'un des sujets ont été exclues en raison d'un défaut d'équipement. Les 29 autres personnes étaient âgées de 19 à 65 ans (âge moyen : 32.52 ans \pm 13.30), et 10 étaient des hommes. Tous ont déclaré avoir une audition normale, et ne présentaient aucun trouble neurologique ou psychologique. Le Comité pour la Protection des Personnes de Besançon a donné son approbation éthique à cette étude sous le numéro 12/653, et le consentement éclairé a été obtenu par écrit de tous les participants.

6.2.3/ ACQUISITION DES DONNÉES

Les participants ont été assis sur une chaise confortable dans une salle d'expérimentation calme, avec pour instruction d'écouter passivement les quatuors de mots qui leur auront

a) Mots congruents



b) Mots incongruents

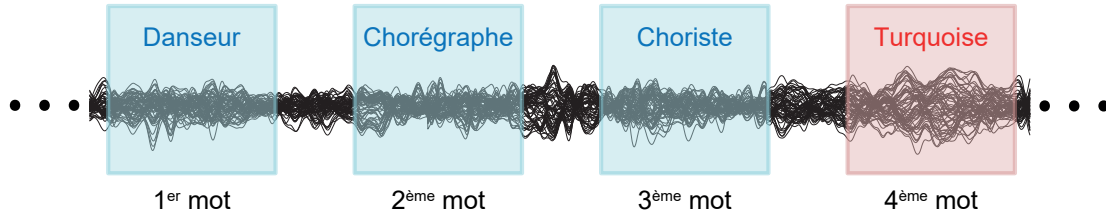


FIGURE 6.1 – Exemple de deux quatuors de mots. a) Le quatrième mot est congruent au contexte sémantique fixé par les trois premiers, ici les arts musicaux. b) Le quatrième mot est incongruent au contexte sémantique : c'est une couleur.

été présentés, sans bouger, et en gardant les yeux fermés. Le logiciel Presentation® (Neurobehavioral Systems, Inc., Berkeley, CA, www.neurobs.com) a été utilisé pour délivrer les stimuli via un casque audio. Le volume sonore a été ajusté individuellement de façon à être confortable. Les données EEG ont été acquises à partir d'un bonnet EEG à 64 canaux suivant le système standard 10/10 (Easycap, EasycapGmbH, Ammersee, Allemagne), à une fréquence d'échantillonnage de 1000 Hz, pour un seuil d'impédance fixé à 20 kΩ. Tous les canaux ont été enregistrés à l'aide de l'équipement numérique OSG (BrainRT, OSG bvba, Rumst, Belgique) avec deux amplificateurs Schwarzer AHNS epas à 44 canaux (Natus, Munich, Allemagne). La durée de l'enregistrement proprement dit était de 20 minutes environ.

6.2.4/ PRÉ-TRAITEMENT

Les données ainsi acquises ont été filtrées dans le logiciel Cartool (version 3.55 ; <http://brainmapping.unige.ch/Cartool.php>) à l'aide d'un filtre passe-bande Butterworth d'ordre 2, entre 1 et 30 Hz et un filtre coupe-bande en encoche a été appliqué afin d'éliminer les artefacts dus à l'alimentation électrique à la fréquence de 50 Hz.

Nous avons ensuite segmenté les données filtrées de toutes les électrodes pour construire des epochs d'une longueur de 900 ms verrouillés temporellement sur le début de chaque stimulus, chaque epoch allant de -100 ms avant le stimulus à 800 ms après le stimulus. Les epochs associés à des électrodes dont la tension dépassait 100 µV ont été éliminés. Les électrodes présentant un signal bruité ont été interpolées à l'aide d'un algorithme de spline tridimensionnel pour l'un des participants (7 électrodes interpolées (sujet 18)). Les epochs ont enfin été corrigés par rapport à la ligne de base en utilisant les 100 ms de signal pré-stimulus.

Une moyenne de 47 epochs par condition (± 2) a été conservée pour chaque sujet.

6.2.5/ SEGMENTATION

Dans l'optique d'examiner la topographie et la latence des ERPs sans *a priori* sur la grande moyenne de groupe, une segmentation en micro-états a été effectuée afin de diviser le signal en périodes de temps pendant lesquelles l'activité électrique est stable en termes de configuration spatiale, en se basant sur le GFP dans les deux conditions, avant de changer brutalement pour une autre cartographie [70, 72]. Un test t de Student apparié a en outre été conduit sur les grandes moyennes individuelles afin de déterminer les régions et les périodes de temps susceptibles de présenter une différence significative ($p < 0.05$) entre les deux conditions congruence / non congruence.

6.2.6/ ANALYSE TEMPS-FRÉQUENCE

Afin d'évaluer la synchronisation à long terme des ondes cérébrales dans le domaine fréquentiel, une analyse temps-fréquence a été menée sur l'ensemble de la longueur du GFP de la grande moyenne, de 1 ms à 800 ms post-stimulus. Cinquante fréquences espacées en échelle logarithmique ont été calculées entre 3 et 30 Hz par transformée de Fourier à fenêtre glissante de Hamming, avec un pas de 10 ms et un ratio de "zero-padding" fixé à 4 [379].

6.2.7/ EXTRACTION ET OPTIMISATION DES CARACTÉRISTIQUES

L'extraction de caractéristiques du signal EEG permet au modèle à développer de se concentrer sur les caractéristiques les mieux corrélées avec le résultat souhaité. Trop de données entrantes sans qualités prédictives produit du bruit et réduit la capacité du modèle à obtenir de bons résultats.

6.2.7.1/ CARACTÉRISTIQUES DU SIGNAL TEMPOREL

Dans un premier temps, nous avons effectué la classification SVM sur les signaux EEG du domaine temporel, sur les paramètres spatio-temporels ressortant comme significatifs lors des tests de Student sur les grandes moyennes de l'ensemble des sujets.

6.2.7.2/ TRANSFORMÉE DE FOURIER À FENÊTRE GLISSANTE

Dans leur format brut, les signaux EEG sont enregistrés dans le domaine temporel. Ainsi, quand on trace un EEG on obtient une représentation temps-amplitude de ce signal. Mais cette représentation n'est pas toujours la meilleure. Dans de nombreux cas, l'information la plus intéressante est cachée dans le contenu fréquentiel du signal. En appliquant une transformée de Fourier (FFT) à un signal brut, on obtient la représentation fréquence-amplitude. Toutefois, si la FFT informe sur la puissance de chaque fréquence présente dans le signal, elle ne dit pas à quel moment cette fréquence existe. Cette information n'est pas nécessaire lorsque le signal est stationnaire, c'est-à-dire si son contenu fréquentiel ne change pas au cours du temps (toutes les fréquences du signal existent à tout moment). Cependant, l'EEG, comme presque tous les signaux biologiques, n'est

pas stationnaire. Pour obtenir la localisation temporelle des composants spectraux, une transformée donnant la représentation temps-fréquence du signal est pertinente.

Nous partirons ainsi du principe que le signal peut être divisé en segments suffisamment petits où il sera considéré comme stationnaire. Dans cet objectif, une fonction fenêtre de Hamming est choisie. Cette fenêtre est au départ localisée tout au début du signal. On suppose que la largeur de la fenêtre est de T ms, et aura un chevauchement de X ms. La fonction fenêtre et le signal sont alors multipliés entre eux. Par la suite, ce produit est considéré comme un autre signal, dont la FFT est calculée. Le résultat de cette transformation est donc la FFT des T premières millisecondes, et si le signal était bien stationnaire, c'est une vraie représentation fréquentielle de ces T premières millisecondes. L'étape suivante consiste à décaler la fenêtre (de t_1 ms), multiplier par le signal et calculer la FFT, et ainsi de suite jusqu'à la fin du signal. En résumé, nous calculons la FFT du signal multipliée par une fonction fenêtre glissante (Figure 6.2).

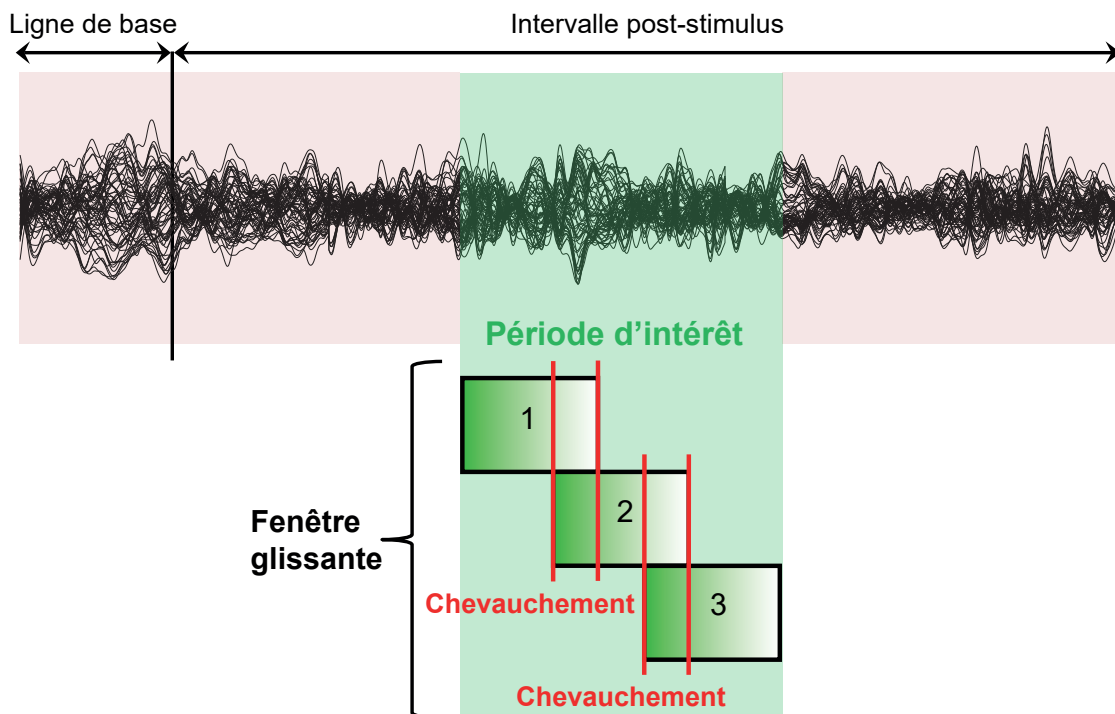


FIGURE 6.2 – Extraction des caractéristiques fréquentielles du signal EEG par FFT à fenêtre glissante. La fenêtre se déplace tout au long de la période d'intérêt sélectionnée dans l'essai EEG, avec un certain chevauchement entre deux fenêtres. Cette méthode permet d'avoir une représentation temps-fréquence et non uniquement fréquentielle du signal.

Le choix d'une bonne fenêtre est particulièrement difficile, et dépend des données : si les composants fréquentiels sont bien séparés les uns des autres dans le signal d'origine, alors on peut sacrifier un peu de résolution fréquentielle pour une meilleure précision temporelle. Mais si ce n'est pas le cas, il sera alors très difficile de trouver une fenêtre optimale. La largeur de l'intervalle de temps considéré fait varier la fréquence minimale observable, pour une fréquence d'échantillonnage fixe. En d'autres mots, plus la fenêtre choisie sera longue dans le temps, plus la fréquence minimale observable sera basse : il est souhaitable d'avoir au moins deux cycles pour extraire une fréquence donnée ; or, plus la fréquence est basse, plus sa période est longue.

Il est donc nécessaire, pour une bonne classification, de déterminer le début et la fin de la période de temps d'intérêt pour l'étude, afin d'éliminer les données inutiles présentes avant et après. D'autre part, le choix de la longueur de la fenêtre glissante pour la FFT et le choix de la longueur du chevauchement seront cruciaux pour extraire les caractéristiques fréquentielles du signal de façon convenable.

6.2.7.3/ UNE OPTIMISATION INDIVIDUELLE

La plupart des études se basent sur la littérature pour déterminer quelles électrodes et quelles périodes de temps sont les plus pertinentes pour obtenir la meilleure *accuracy* de classification. La N400 est un potentiel négatif dont le pic se présente environ 400 ms post-stimulus, observable entre environ 250 et 550 ms. Ainsi, la période de temps allant de 250 à 550 ms post-stimulus semble la plus pertinente pour la classification, et ce pour tous les sujets.

Cependant, il existe dans l'activité cérébrale de grandes variations d'un sujet à l'autre, y compris chez les sujets sains, comprenant des fluctuations en latence mais également en topographie. Nous avons donc estimé qu'il était nécessaire de personnaliser l'extraction de ces paramètres afin de s'adapter au profil de chaque sujet testé.

Ainsi, nous allons utiliser un algorithme dit "génétique" pour rechercher les valeurs optimales de chacun de ces quatre paramètres : le début et la fin de la période d'intérêt, sur laquelle sera effectuée la FFT, la longueur de la fenêtre glissante, et la longueur du chevauchement lors du déplacement de la fenêtre glissante (Figure 6.3).

6.2.7.4/ ALGORITHME GÉNÉTIQUE

Les algorithmes génétiques sont une famille de méthodes de calcul de recherche aléatoire basée sur les principes de la génétique naturelle et de l'évolution. Ils permettent de résoudre des problèmes d'optimisation globale en faisant évoluer une population de solutions candidates.

Chaque individu de la population de solutions potentielles est représenté par un vecteur appelé "chromosome", constitué d'un ensemble de caractéristiques à optimiser, les "gènes". Ces caractéristiques à optimiser sont ici le début et la fin de la période d'intérêt, la longueur de la fenêtre glissante de la FFT et le chevauchement entre deux fenêtres glissantes. La première population est initialisée en remplissant de façon aléatoire les vecteurs de caractéristiques. Une fonction d'aptitude du modèle est utilisée pour évaluer chaque chromosome de la population. Nous considérons le taux d'erreur du classificateur comme la fonction d'aptitude que le GA s'efforcera de minimiser. Par conséquent, pour chaque chromosome, un SVM est créé, entraîné et testé pour obtenir ce taux d'erreur. À chaque étape, le GA créera des "enfants" destinés à faire partie de la prochaine génération de la population. Toutefois, la probabilité qu'un chromosome particulier soit choisi pour créer la prochaine génération dépend des performances obtenues lors de la classification.

Trois types d'enfants peuvent être créés : les enfants élite, qui ont le plus faible taux d'erreur pour cette génération, survivent automatiquement pour la prochaine. Cela réduit le risque de perdre les meilleurs chromosomes par croisement ou mutation et permet une convergence plus rapide vers une solution optimale. Parmi les chromosomes restants, les

meilleurs sont sélectionnés pour être parents et se recombinent, donnant naissance à des descendants potentiellement supérieurs, les enfants issus de croisement. Pour chaque couple de parents, des gènes (valeurs des paramètres à optimiser) sont choisis au hasard dans l'un ou l'autre et attribués au chromosome de l'enfant. Un même individu peut être sélectionné plus d'une fois pour être parent, auquel cas ses gènes contribuent à créer plus d'un enfant. Enfin, des mutations peuvent être introduites en changeant au hasard une valeur dans un chromosome pour créer un enfant muté. Ces mutations contribuent à la diversité de la population, et augmentent la probabilité que le GA génère des individus capables d'atteindre un meilleur taux d'erreur. En outre, ces modifications permettent également à l'algorithme d'explorer de nouveaux domaines du problème.

L'ensemble du processus est répété jusqu'à ce que la variation moyenne du meilleur taux d'erreur obtenu soit inférieure à la tolérance fixée à 10^{-6} , ou que l'on atteigne un taux d'erreur de zéro, soit une discrimination parfaite. Après la dernière itération, le meilleur chromosome, c'est-à-dire les valeurs des paramètres pour la FFT offrant la meilleure performance de classification, est retenu et considéré comme la solution optimale pour ce problème.

Ce processus d'optimisation a été répété pour chacun des 29 sujets sains et 5 patients afin d'obtenir les meilleures paramètres de façon personnalisée.

6.2.8/ SUPPORT VECTOR MACHINE

Le SVM implique de déterminer, lors de la phase d'entraînement l'hyperplan optimisant la distance aux points les plus proches de chaque classe, séparant au mieux les données. Lors de la phase subséquente de test, le reste des essais est présenté au modèle appris, qui les classera alors selon les paramètres définis précédemment.

Une validation croisée "*k-fold*" a été effectuée sur ce SVM : les essais ont été randomisés et divisés en quatre groupes, chacun d'entre eux servant d'échantillon de test pendant que les trois autres constituaient l'échantillon d'entraînement. Cette opération a été répétée 1000 fois, et un pourcentage de discrimination (*accuracy*) ainsi qu'une valeur de p-value ont alors été calculés. Dans un premier temps, le résultat obtenu à l'issue de ce procédé imbriqué dans l'algorithme génétique a été considéré.

Dans un second temps, 10 essais ont été mis de côté et ne sont pas passés par le processus d'optimisation par l'algorithme génétique. A l'issue de l'optimisation des paramètres de la FFT, ces 10 essais tenus secrets pour l'algorithme de recherche ont été testés sur le modèle optimisé obtenu, afin de comparer le résultat obtenu lorsque l'algorithme avait eu accès à l'ensemble des données ou non.

6.3/ RÉSULTATS

La segmentation effectuée sur les ERPs a révélé six micro-états de topographie cérébrale stable (Figure 6.4). Les deux premiers présentent nettement des ERPs dans les deux conditions congruence et non-congruence. Le premier micro-état, allant de 16 à 74 ms post-stimulus présente une distribution fronto-centrale positive pour la condition "mots non congruents" et davantage centrale pour la condition "mots congruents" ($p < 0.001$). Le premier intervalle, courant de 74 à 149 ms, montre une carte avec une forte négativité

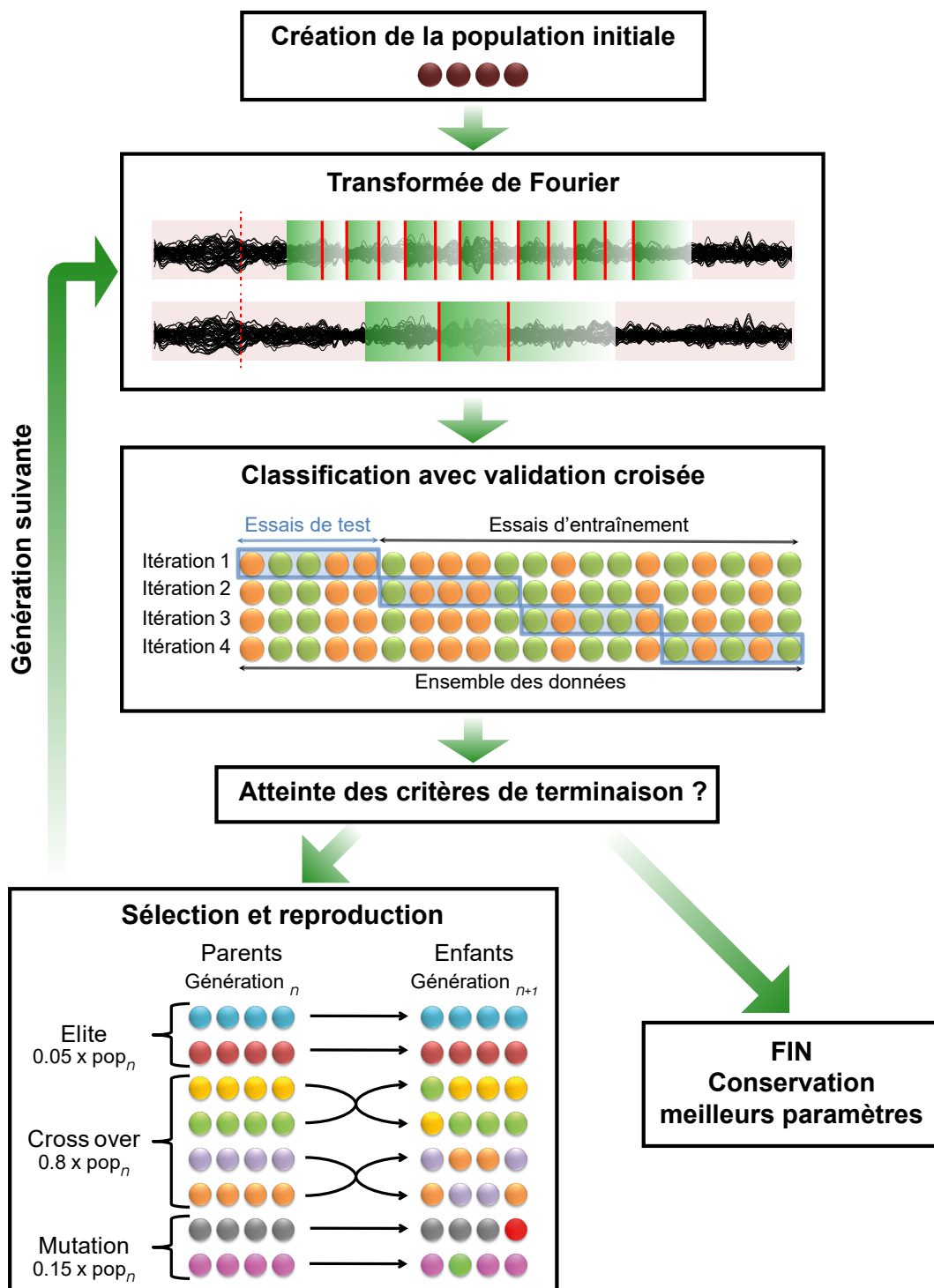


FIGURE 6.3 – Pipeline incluant l'extraction des caractéristiques et la classification des essais au sein de l'algorithme d'optimisation génétique pour chacun des sujets.

centro-pariétale, puis une composante positive centrale a pu être observée dans le second micro-état, entre 150 et 253 ms, et une troisième onde, négative et centro-pariétale était présente de 254 à 406 ms. Aucune différence significative n'a pu être mise en évidence entre les deux conditions pour ces trois cartes. Les micro-états quatre et cinq montrent une différence significative ($p < 0.001$) dans la région centro-pariétale. En effet,

entre 407 et 505 ms, on observe une onde négative fronto-centrale évoluant vers une forme frontale entre 506 et 663 ms lorsque les mots sont congruents, tandis que cette négativité est débute dans la zone centro-pariétale avant de devenir fronto-centrale dans la seconde condition. Enfin, la sixième cartographie, de 664 à 800 ms, est à nouveau commune aux deux conditions, avec un complexe positif en centro-pariétal et négatif dans la région frontale.

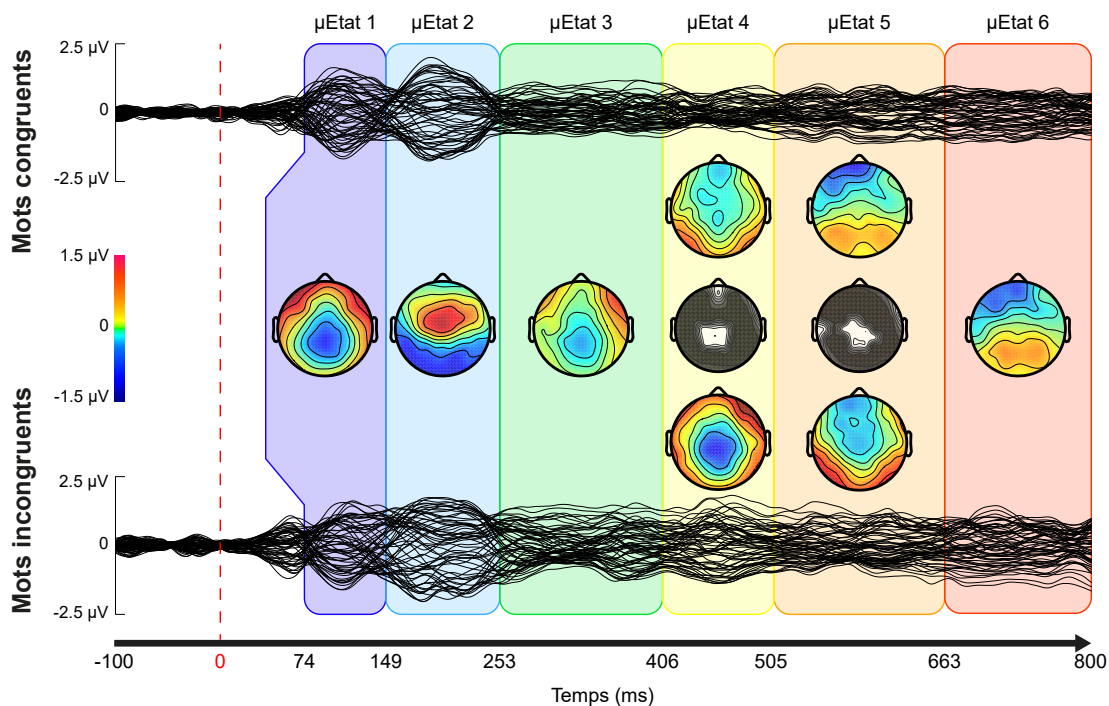


FIGURE 6.4 – Grandes moyennes de l'ensemble des sujets pour les conditions congruente et incongruente et cartographies du signal des 63 électrodes. Six topographies ont été établies après la réalisation de la segmentation, montrant l'émergence de différentes périodes de stabilité, correspondant à des mécanismes cognitifs particuliers à un moment donné.

Une analyse de la densité spectrale de puissance a été exécutée afin d'examiner les topographies sur le long terme sur 4 différentes fréquences situées respectivement dans les bandes thêta, alpha, bêta 1 et bêta 2 (Figure 6.5). Aucune différence significative n'a pu être mise en évidence entre les mots congruents et incongruents dans ces cartographies.

L'analyse temps-fréquence du GFP de la grande moyenne (Figure 6.6) révèle une synchronisation plus forte de l'activité cérébrale, entre 5 et 15 dB, dans les bandes de fréquence thêta, alpha et bêta 1 pour les mots congruents dans une première période allant de 150 à 400 ms après le début du stimulus, ainsi que dans les bandes bêta 1 et bêta 2 entre 430 et 620 ms. Deux périodes distinctes montrent en revanche une synchronisation plus forte pour les mots incongruents, dans les bandes thêta et alpha de 560 à 620 ms, et dans les fréquences alpha et bêta 1 entre 800 et 900 ms.

La figure 6.7 montre que l'utilisation des paramètres issus du test de Student effectué dans le domaine temporel suite à la segmentation en micro-états donne des précisions de test proches de l'aléatoire, avec une moyenne de $49.5\% \pm 4.2\%$ sur l'ensemble des

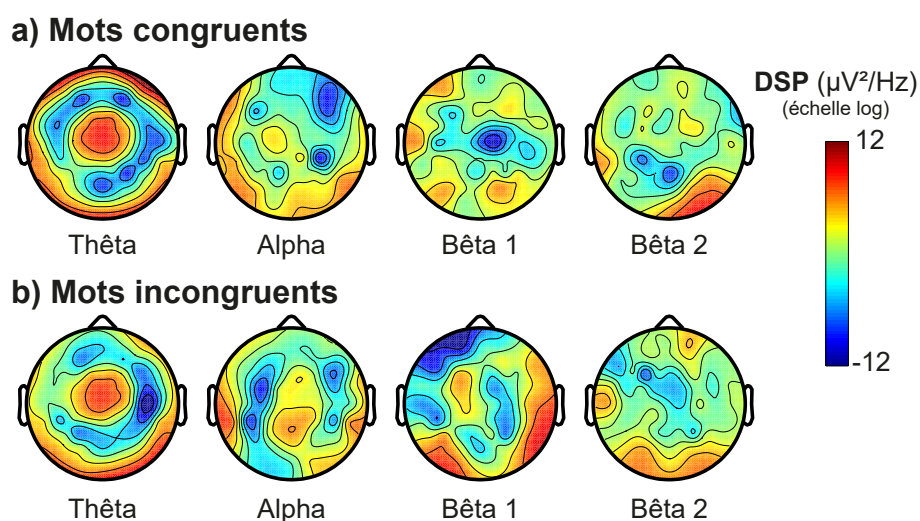


FIGURE 6.5 – Cartographies du signal EEG dans le domaine fréquentiel, obtenues sur les grandes moyennes de l’ensemble des sujets, pour chacun des bandes de fréquences considérées et pour a) la condition congruente et b) la condition incongruente.

sujets. La précision maximale obtenue est de 58.4%, pour le sujet 8.

La même figure (Figure 6.7) montre que la performance en termes de pourcentage de précision lors de l’utilisation de paramètres “standards” d’extraction de caractéristiques basés sur la littérature est de 48.47% en moyenne ($\pm 5.20\%$). En revanche, la précision après optimisation individuelle des paramètres via l’algorithme génétique est systématiquement plus élevée que la précision standard pour tous les sujets, avec une moyenne de 67.70% ($\pm 3.74\%$). Il est intéressant de constater que la performance varie fortement d’un sujet à l’autre lorsque ces paramètres sont testés sur des essais mis de côté et inconnus de l’algorithme d’optimisation. En effet, le processus a conduit à une amélioration de la performance chez 14 sujets, tandis que pour 7 d’entre eux aucune variation significative de la précision n’a eu lieu, et que pour 10 sujets enfin, la performance s’est même dégradée (moyenne : 49.14% \pm 15.72).

Le nombre de sujets dont la période d’intérêt optimale inclut un temps donné est figuré par la ligne orange de la Figure 6.8. La proportion de sujets excédant 75% concerne une fenêtre de temps comprise entre 312 et 563 ms après la survenue du stimulus. Toutefois, on peut constater de grandes disparités entre les fenêtres optimales sélectionnées pour chacun des sujets, tant en termes de longueur (moyenne : 448 ± 145 ms), que du temps de début (moyenne : 252 ± 124 ms) et de fin (moyenne : 700 ± 145 ms).

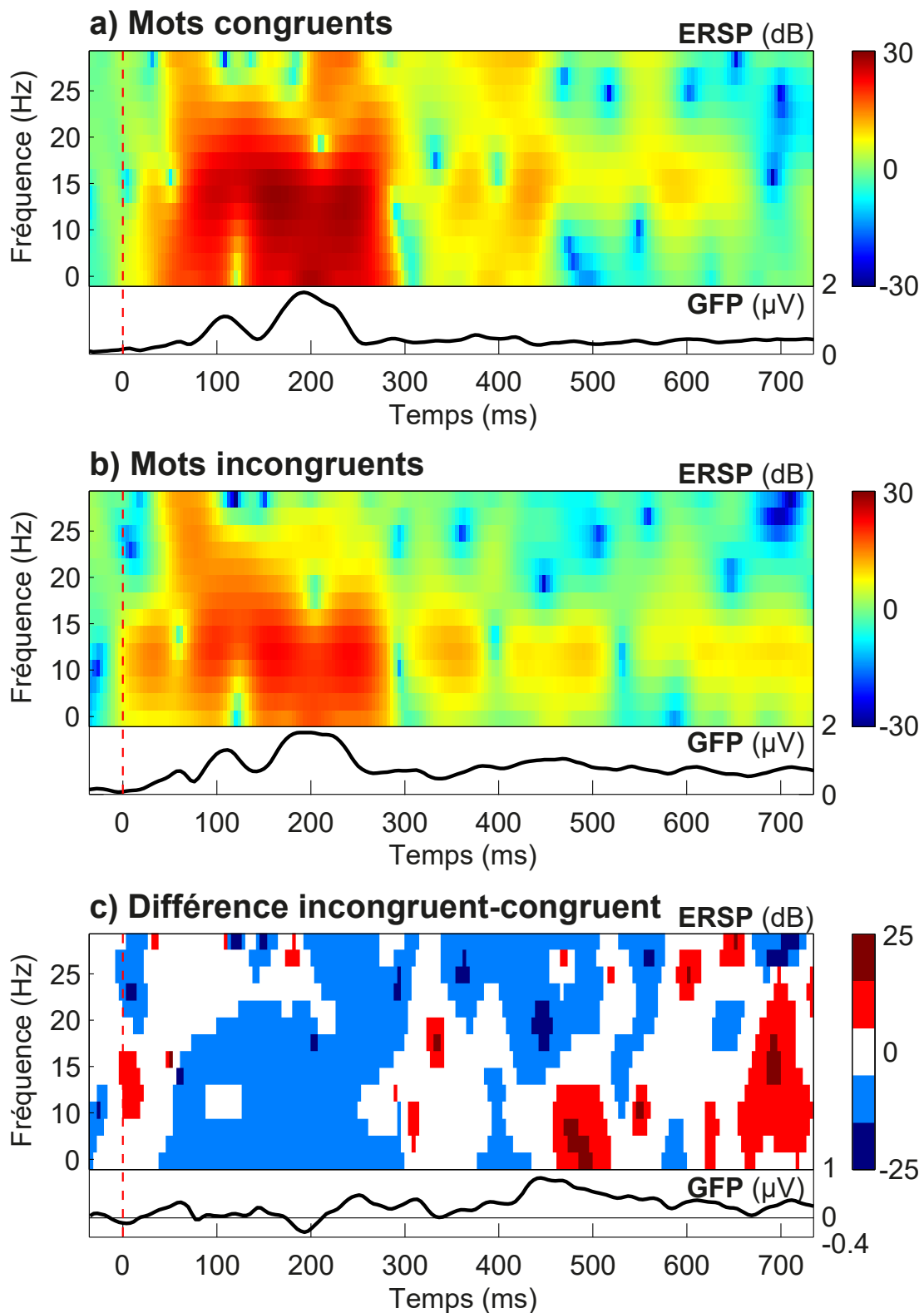


FIGURE 6.6 – Analyse temps-fréquence sur le GFP de chacune des deux conditions (a) mots congruents et b) mots incongruents). Une forte synchronisation (rouge) suivie d'une désynchronisation (jaune et bleu) est clairement observée pour chaque catégorie. Une différence dans la synchronisation entre les deux conditions est observée en c).

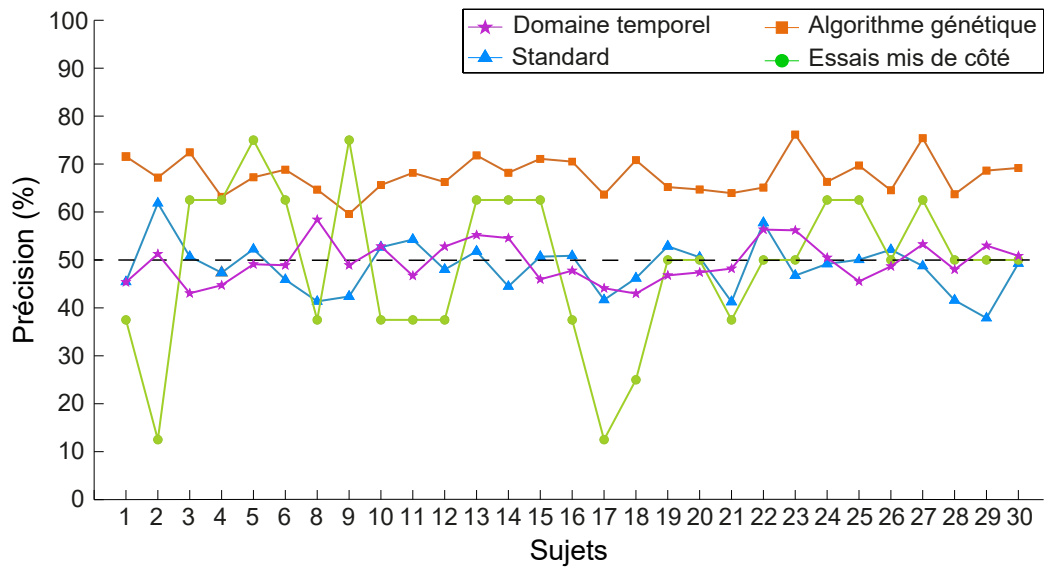


FIGURE 6.7 – Précision de test obtenue pour chacun des sujets avec une méthode de classification utilisant :

- dans le domaine temporel : les électrodes (C3, Cz, CP5, CP1, P3, Pz, C1, C2, CP3, CPz, P1, P2) et la période d'intérêt (de 407 à 653 ms) significatives lors de la segmentation (étoile violette),
- dans le domaine fréquentiel : des paramètres de FFT standards, identiques pour tous les sujets (triangle bleu), précision de validation obtenue par optimisation avec algorithme génétique (carré orange), précision de test obtenue sur les essais mis de côté et inconnus de l'algorithme génétique afin de vérifier la performance (rond vert).

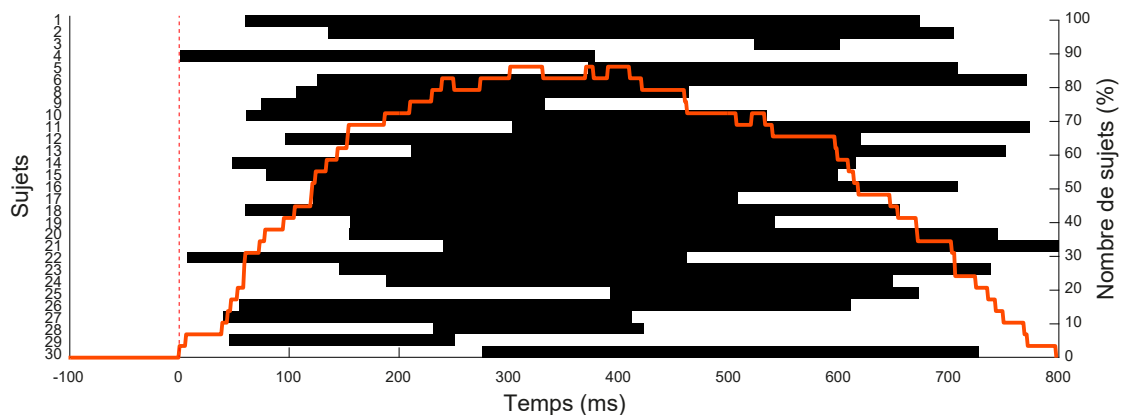


FIGURE 6.8 – Position de la période d'intérêt optimale retenue pour chaque sujet par l'algorithme génétique. La courbe orange représente le pourcentage de sujets pour lesquels chaque milliseconde fait partie de la période optimale.

La Figure 6.9 donne pour chaque sujet, puis sous forme d'histogrammes, les longueurs respectives de la période d'intérêt (en vert), de la fenêtre glissante (en bleu) et du chevauchement (en orange) optimaux obtenus par algorithme génétique. Là encore, une forte dispersion des valeurs peut être observée d'un participant à l'autre concernant les pa-

ramètres de la transformée de Fourier, suggérant que l'emploi de paramètres standards communs à tous les sujets ne serait pas judicieux.

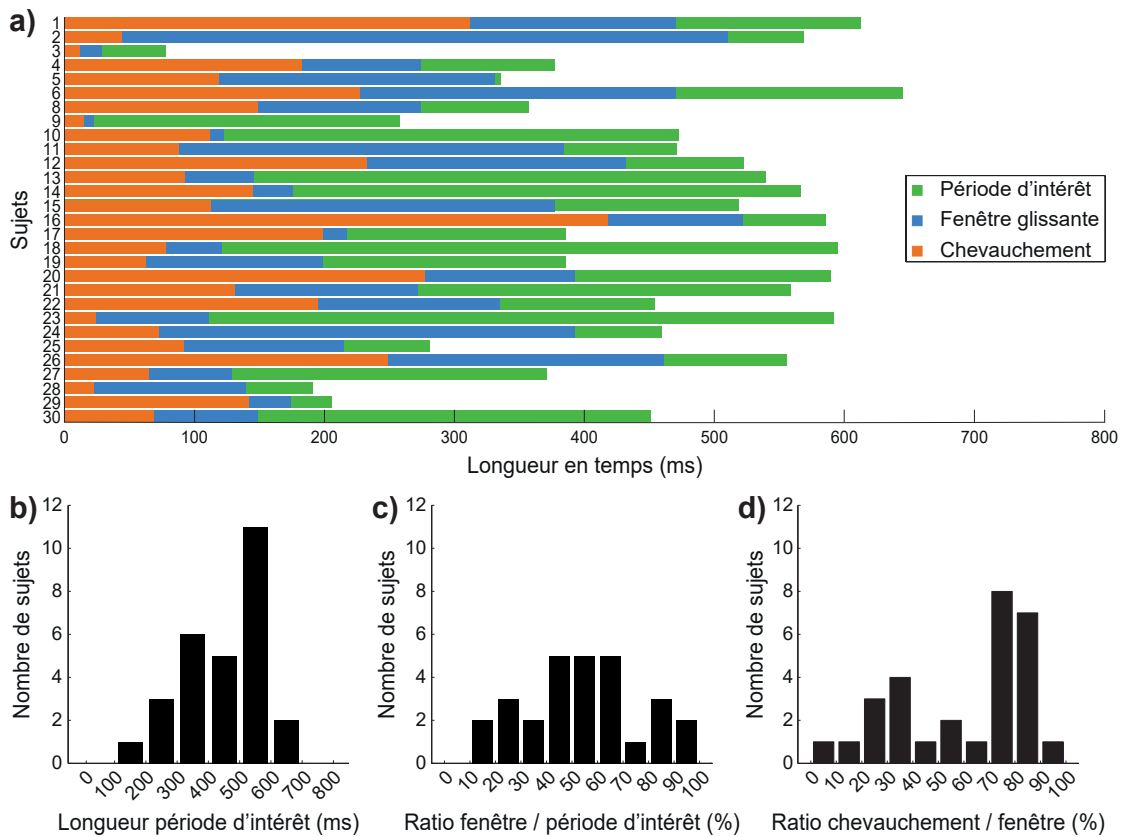


FIGURE 6.9 – Paramètres optimaux obtenus par algorithme génétique. a) Longueurs en temps de la période d'intérêt, de la fenêtre glissante et du chevauchement optimaux pour chaque sujet. b) Nombre de sujets par longueur optimale de la période d'intérêt. c) Nombre de sujets par ratio longueur optimale de la fenêtre glissante sur longueur optimale de la période d'intérêt. d) Nombre de sujets par longueur optimale du chevauchement sur longueur optimale de la fenêtre glissante.

Un test statistique de Spearman (ρ) a été opéré sur les différents résultats obtenus. Plus la corrélation est forte et positive, plus le coefficient est proche de 1, plus la corrélation est forte et négative, plus le coefficient est proche de -1, et lorsque la tendance n'est pas monotone, il est proche de zéro. La Table 6.1 démontre qu'aucune corrélation ne peut être faite entre les paramètres déterminés comme optimaux par l'algorithme génétique et la performance obtenue tant par l'algorithme qu'en test sur les essais inconscients de celui-ci par la suite.

La Figure 6.10 montre l'évolution de la précision obtenue lors de l'application du processus d'optimisation par algorithme génétique à des données aléatoires, en fonction du nombre d'essais par classe. On constate qu'en deçà de 300 essais par classe, la performance est inversement proportionnelle au nombre d'essais par classe, suggérant que l'algorithme génétique est capable de distinguer deux classes d'après des motifs aléatoires lorsque le nombre d'essais est trop faible. Il serait donc recommandé d'enregistrer un nombre important d'essais lors des études afin de se prémunir contre cet effet.

	Size of period of interest	Ratio size sliding window/period of interest (%)	Ratio size overlap/sliding window (%)	Précision algorithme génétique	Précision de test
Size of period of interest					
Ratio size sliding window/period of interest (%)	-0.06				
Ratio size overlap/sliding window (%)	0.13	-0.22			
Précision algorithme génétique	0.12	0.11	0.36		
Précision de test	-0.02	0.06	-0.26	-0.02	

Tableau 6.1 – Corrélations de Spearman (ρ) entre les différents paramètres optimisés et la précision de test obtenue sur les essais préservés.

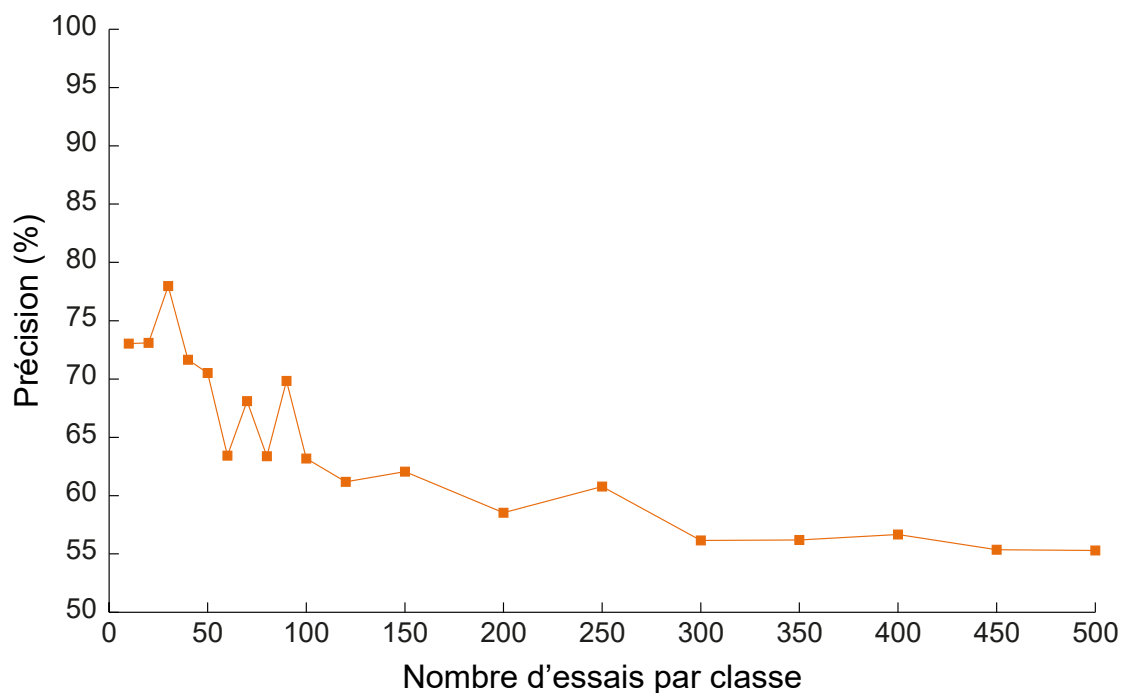


FIGURE 6.10 – Optimisation sur données générées aléatoirement. Précision de validation obtenue par optimisation par algorithme génétique en fonction du nombre d'essais par classe.

6.4/ DISCUSSION

La présente investigation avait pour objectif d'examiner la possibilité de classer automatiquement et de façon individualisée les potentiels évoqués liés au traitement sémantique. Nous avons opposé des quatuors de mots dont le dernier élément était ou non relié au contexte sémantique mis en place par les trois premiers, en formulant l'hypothèse que l'institution dudit contexte via des quatuors de mots plutôt que par des paires induirait une différence marquée en électrophysiologie entre les deux conditions, et donc une meilleure performance de classification. Notre seconde hypothèse était que l'optimisation individualisée de l'extraction des caractéristiques utilisées pour le SVM améliorerait nettement la précision de la classification.

La N400 est un potentiel évoqué négatif, apparaissant dans les régions centro-pariétales du scalp, dont le pic est atteint environ 400 ms après le début du stimulus. L'amplitude de cette onde serait spécifiquement sensible à la composante sémantique de l'information auditive reçue. L'intérêt majeur de cet ERP dans notre étude est que l'observation d'un effet N400 chez le sujet reflète un accès au sens du discours, dont la préservation serait corrélée à un diagnostic favorable chez le patient en état de conscience altérée.

Dans cette étude, nous observons une série d'activations cérébrales lors de la présentation du quatrième mot du quatuor. Ces activations comprennent tant des réponses attentionnelles que cognitives, dont la N400 recherchée. Nos analyses ont révélé que l'amplitude et la durée de ce potentiel évoqué sont bien sensibles à l'incongruence au contexte du quatrième mot. L'utilisation de l'algorithme génétique couple au SVM permet une augmentation drastique de la performance de classification, mais la précision obtenue en test n'étant pas en rapport, nous avons fait l'hypothèse que l'optimisation pouvait être parasitée par le faible ratio signal sur bruit des enregistrements EEG, menant à un phénomène de surajustement.

6.4.1/ POTENTIELS ÉVOQUÉS

La segmentation du signal EEG nous a permis d'étudier l'activité transitoire correspondant au traitement à court terme de l'information sémantique. Nous avons mis en évidence l'apparition successive de 6 périodes d'intérêt après l'apparition du stimulus.

6.4.1.1/ RÉPONSE CÉRÉBRALE PRÉCOCE

Aucune différence significative n'a pu être mise en évidence au niveau de la première cartographie, courant de 74 à 149 ms. Nous supposons que cette topographie à la négativité centro-pariétale correspond à une onde N100, comme le suggère le pic d'amplitude à environ 100 ms. L'amplitude de la N100, élicitée par des stimuli auditifs discernables, a été interprétée comme reflétant l'allocation de l'attention sensorielle auditive de façon spontanée [380].

6.4.1.2/ MÉCANISMES ATTENTIONNELS

La seconde cartographie (150 - 253 ms) ne présente pas non plus de différence significative entre les deux conditions. Le maximum d'amplitude pouvant être observé à environ

200 ms et la distribution fronto-centrale positive pointent vers une onde P200, supposée représenter le traitement perceptuel secondaire des stimuli. L'amplitude est proportionnelle à la quantité de ressources attentionnelles consacrées à la tâche et au degré de traitement de l'information requis [381].

Présente de 254 à 406 ms, la troisième topographie représente une déflexion négative en centro-pariétale semblable à celle de la N100, mais avec une amplitude bien plus faible. Cette carte est plus difficile à interpréter, car nous nous serions attendus à observer plutôt un pic positif de P300 à ce stade. Cette distribution, qui ne montre pas de différence significative entre les deux conditions, ressemble à une période de transition entre la P200, avec une positivité antérieure forte et une négativité postérieure également forte, vers une négativité frontale et une positivité pariéto-occipitale élevées, pour les mots congruents, observables sur les micro-états 6 et 7 ; ou vers une négativité centro-pariétale forte présente dès le micro-état 4 dans le cas des mots incongruents.

6.4.1.3/ RÉPONSES COGNITIVES

Comparé à la condition congruente, les mots non-congruents ont clairement élicité un potentiel négatif centro-pariétal reflétant l'incongruité sémantique présente dans cette circonstance [382] entre 406 et 505 ms. Il semble que l'utilisation de quatuors plutôt que de paires de mots induise une attente forte en termes de contexte, induisant une N400 fiable.

La question de savoir si la N400 est automatique ou contrôlée par l'attention est sujette à débat, (voir par exemple [383–385]). Toutefois, ici nous n'observons pas de différence dans la réallocation préalable de l'attention dans les processus attentionnels précédant la survenue de la N400.

Le dernier micro-état, couvrant la période de temps 663 - 800 ms, montre une négativité frontale et une positivité pariéto-occipitale fortes, la topographie pouvant correspondre à la *Late Positive Component* (LPC). Il est probable que la LPC reflète ici l'intégration sémantique et la confiance dans l'intégration d'un mot dans son contexte [386]. Il a été suggéré [387, 388] que la LPC dépend d'un rappel en mémoire conscient [389] de la relation entre la cible et l'amorce, ici entre le quatrième mot et le contexte précédemment établi.

Il est intéressant de constater que la dernière cartographie ne montre pas de différence significative entre les deux conditions. Il semble que dans la condition congruente, une phase de transition précède une arrivée de la LPC dès la cinquième carte et se prolongeant sur la dernière, tandis que le décours temporel en condition incongruente distingue d'abord une N400 sur les quatrième et cinquième cartes, avant l'apparition enfin de la LPC sur le dernier micro-état, provoquant cette différence significative ($p < 0.001$) dans la zone centro-pariétale.

6.4.2/ ACTIVITÉ FRÉQUENTIELLE : CONSÉQUENCES POUR LA CLASSIFICATION

L'examen de l'activité fréquentielle dans les bandes Thêta, Alpha, Bêta 1 et Bêta 2 sur l'ensemble de l'époque, soit de 0 à 800 ms après le début du quatrième mot, n'a révélé aucune région d'intérêt significative pour distinguer les deux conditions. Nous avons donc choisi de conserver l'ensemble des électrodes pour la classification, afin de ne pas perdre

d'information pouvant être pertinente.

En revanche, l'analyse temps fréquence a révélé des fenêtres de temps intéressantes, pouvant concerner l'ensemble des quatre bandes de fréquence. Il nous a donc semblé pertinent, dans un premier temps, d'opérer la classification sur une fenêtre de temps typique de la N400 et susceptible de contenir une différence significative. Dans un second temps, au vu des résultats obtenus et connaissant les variations inter-individuelles susceptibles de survenir en terme de latence de l'activité cérébrale, nous avons choisi d'optimiser la fenêtre de temps de façon personnalisée pour chacun des sujets.

6.4.3/ UNE FORTE VARIABILITÉ INTER-INDIVIDUELLE

L'ensemble des résultats obtenus lors de la classification pointe vers la confirmation d'une forte variabilité inter-individuelle, comme en témoignent plus particulièrement les fluctuations de période d'intérêt, de taille de celle-ci, ainsi que de longueur de la fenêtre glissante et du chevauchement d'un sujet à l'autre. Il semblerait donc que l'utilisation d'une optimisation personnalisée soit pertinente dans cette étude. En outre, le fait qu'aucune corrélation n'ait pu être établie entre la performance et les valeurs obtenues pour les paramètres optimisés tend à confirmer que ceux-ci dépendent essentiellement du sujet en question.

6.4.4/ UN SURAJUSTEMENT DU MODÈLE

6.4.4.1/ SURAJUSTEMENT DÛ À L'OPTIMISATION DES PARAMÈTRES

La méthode de classification avec EEG-HR que nous présentons ici prend en compte toutes les régions, pas seulement celles où la N400 est supposée apparaître, et la période de temps est optimisée individuellement, rendant cette façon de faire particulièrement adaptée pour les patients souffrant de lésions cérébrales, dont les potentiels évoqués peuvent être répartis selon d'autres topographies que les sujets sains dues aux dysfonctionnements cérébraux des patients [361, 362].

Toutefois, conserver l'ensemble des 63 électrodes signifie inévitablement introduire davantage de données non pertinentes pour éviter de manquer celles qui le sont. La classification prend en compte toutes les données entrées dans le SVM, y compris celles qui ne sont pas pertinentes pour différencier les deux classes. Ainsi, d'autres processus cérébraux que la N400, ou du bruit lié à l'enregistrement, peuvent avoir contribué à la création du modèle. Le nombre d'essais dont nous disposons est très faible, l'influence du bruit augmente donc. Or, ainsi que le démontre la Figure 6.10, lorsque l'on augmente le nombre d'essais générés aléatoirement, la précision de validation obtenue via l'algorithme génétique diminue pour se rapprocher des 50% attendus, signifiant que les données sont classées au hasard et qu'aucun pattern dû au bruit n'est responsable de la performance du modèle. Plus la taille du jeu d'entraînement s'amenuise, plus le pourcentage de précision au-delà duquel on peut considérer que la performance n'est pas due au hasard augmente. De plus, si l'utilisation de 120 quatuors de mots différents permet de varier les stimuli et d'éviter ainsi tant l'habituation du sujet que l'anticipation du quatrième mot, nous ne connaissons pas l'influence que peut avoir l'usage de stimuli tous différents. Il est probable que cette méthode introduise une variance supplémentaire

dans l'activité cérébrale, qui s'annule lors de l'utilisation des moyennes de groupe mais qui resurgirait lorsque l'on revient à une classification essai par essai.

Ces différents facteurs nous permettent d'expliquer que l'efficacité du modèle SVM créé varie autant de l'entraînement au test : si des données représentatives de la réalité sont utilisées pour l'entraînement, la précision de test est une alors estimation non biaisée de la réelle capacité de généralisation du SVM.

La généralisation est la capacité d'un modèle à donner de bons résultats sur des essais inconnus. Atteindre une bonne généralisation est l'un des objectifs les plus importants de toute méthode d'apprentissage automatique. On dit qu'il y a surajustement (*overfitting*) lorsqu'un modèle de classification obtient de bons résultats sur les essais d'entraînement mais pas sur les essais de test : il perd sa capacité de généralisation de l'apprentissage et produit des résultats trompeurs. Cela indique que les relations sous-jacentes de l'ensemble des données n'ont pas été apprises, et qu'au lieu de cela, un ensemble de relations existant uniquement sur les données d'entraînement a été appris, mais celles-ci n'ont aucune correspondance sur l'ensemble des cas connus. En d'autres termes, le classificateur créé lors de l'apprentissage s'adapte si bien aux essais d'entraînement que le bruit et les particularités de ce jeu de données sont mémorisés, et que ce modèle est limité à quelques données spécifiques au lieu de données générales [390, 391].

Les caractéristiques particulières des données EEG, c'est-à-dire le petit nombre d'échantillons (moins de 100 par sujet dans notre cas) et le grand nombre de caractéristiques extraites, dû notamment au grand nombre d'électrodes conservées, augmentent donc théoriquement la probabilité d'un surajustement du SVM.

Afin d'obtenir davantage d'essais, les stimulations passives pourraient être divisées en plusieurs sessions. Toutefois, cela ouvre la porte à d'autres défis, puisque l'état de base de l'activité cérébrale et les conditions dans lesquelles se trouve le sujet varient d'une session à l'autre, introduisant davantage de variabilité intra-individuelle.

6.4.4.2/ SURAJUSTEMENT DÛ À DES HYPERPARAMÈTRES DU SVM TROP RIGIDES "STANDARDS"

D'autre part, nous avons fait le choix ici d'utiliser un *kernel* linéaire, avec un hyperparamètre de régulation C standard, fixé à 1. Les hyperparamètres sont très importants pour construire des modèles robustes et précis et donc éviter que le modèle ne soit trop ou pas assez ajusté. Plus C est grand, plus le SVM va tenter d'améliorer la performance à l'entraînement, menant à un surajustement du modèle aux données d'entraînement et donc une performance plus faible lors du test. Dans cette étude, il est possible que le surajustement observé soit en partie dû à un C trop élevé. Toutefois, cette division en paramètres et hyper-paramètres nécessite d'ajuster le modèle de manière itérative en utilisant une paire de boucles imbriquées, les hyperparamètres étant ajustés dans la boucle extérieure (sélection du modèle) et les paramètres étant réglés dans la boucle intérieure (ajustement du modèle), ce qui est très coûteux en ressources computationnelles et très long. D'autre part, d'après Cawley & Talbot (2007) [392], le recours à cette méthode prête le flanc à un surajustement dû à l'optimisation des hyperparamètres, en sus de celui lié à l'optimisation des paramètres d'extraction des caractéristiques.

Ainsi, nous montrons que certains protocoles d'évaluation des performances, en apparence bénins, utilisés couramment par la communauté du *machine learning* en EEG,

sont susceptibles de cette forme de biais et donnent donc potentiellement des résultats erronés.

Afin d'éviter ce biais, l'optimisation des paramètres et hyperparamètres doit être traitée comme une partie intégrante du processus d'ajustement du modèle et doit être renouvelée chaque fois qu'un modèle est ajusté à un nouveau set d'entraînement. Cependant, cela signifie qu'une évaluation robuste et impartiale des performances nécessitera probablement des protocoles plus rigoureux et plus gourmands en calculs, comme la validation croisée emboîtée ou "double croix" [393].

De toute évidence, une mauvaise performance de classification ne doit pas être prise comme une preuve de l'absence de conscience chez des patients en état de conscience altérée, puisque même les potentiels évoqués d'intérêt peuvent rester indétectés même chez des sujets sains.

6.5/ CONCLUSION

Il apparaît extrêmement difficile de définir quelle serait la meilleure méthode d'analyse du signal, tant elle dépend du sujet, du design du protocole, du nombre d'essais disponibles, pour ne citer que quelques uns des nombreux facteurs d'influence. Il semble, de plus, que l'utilisation d'une méthode plus complexe et, ou personnalisée, ne suffise pas à résoudre l'incertitude du diagnostic, voire puisse éventuellement compliquer la prise de décision quant à la prise en charge médicale. Une comparaison systématique des différentes techniques d'analyse et de *machine learning* sur un même jeu de données, puis sur différents jeux de données issus de protocoles différents pourrait permettre de dégager une vision d'ensemble claire des prérequis nécessaires à l'obtention d'une performance de classification à la fois haute et fiable, et devra faire l'objet de futures études.

IV

DISCUSSION GÉNÉRALE ET CONCLUSION

DISCUSSION GÉNÉRALE

7.1/ CONTRIBUTIONS DE CETTE THÈSE

Nous avons débuté ce manuscrit par une revue systématique de la littérature, qui a mis principalement en relief un manque de reproductibilité des études sur le SVM appliqué à l'EEG. Cet état de fait est retrouvé largement dans la littérature depuis quelques années, à tel point que l'on parle maintenant de "crise de la reproductibilité" [394–396].

Dans notre revue, comme de manière plus générale, cette problématique trouve ses sources dans un manque d'accessibilité des données et des programmes d'analyse. Nous avons notamment mis en lumière que 98.6% des articles de notre panorama étaient très difficiles ou impossible à reproduire.

Pour ce qui est du code servant à l'analyse, il devrait être relativement aisé de le mettre en ligne sur des plateformes telles que GitHub, ou en tant que matériel supplémentaire accompagnant la publication de l'article. Dans tous les cas, le logiciel et/ou le langage de programmation, ainsi que les éventuelles boîtes à outils numériques utilisées et leurs versions devraient être explicitement mentionnés dans le corps de texte.

La question des enregistrements EEG est, en revanche, d'une toute autre complexité. En effet, les données en rapport avec la santé de l'humain acquises à l'hôpital, et même dans le cadre de toute recherche, sont protégées par le Règlement Général sur la Protection des Données¹ en Europe, et par des équivalents dans d'autres pays. Il n'est donc pas toujours possible de les mettre à disposition d'autres équipes de recherche. C'est notamment le cas des données exploitées dans ce travail de thèse. Cette situation entraîne deux conséquences :

- L'étude n'est pas reproductible directement par une autre équipe, ce qui nuit à sa portée et à la robustesse de ses conclusions. Toutefois, dans la plupart des cas, les études cliniques de neuroimagerie sont d'abord validées sur sujets sains. Il devrait être possible de partager ces enregistrements anonymes afin que la communauté scientifique puisse entériner la fiabilité des nouveaux protocoles.
- Lorsque l'on veut tester cette fois une nouvelle méthode d'analyse, par exemple en *machine learning*, il est difficile de trouver en ligne des bases de données EEG qui soient adaptées aux problématiques cliniques visées. En effet, la plupart des banques proposent surtout des données de BCI (compétitions BCI), ou en rapport avec les émotions (jeu de données DEAP), qui sont les deux champs les plus ex-

1. Loi n° 78-17 du 6 janvier 1978 relative à l'informatique, aux fichiers et aux libertés, JORF n° 0006 du 7 janvier 1978 ; modifiée par l'ordonnance n° 2018-1125 du 12 décembre 2018.

plorés. Nous soulignons donc ici la nécessité de pouvoir compter sur un catalogue plus vaste pour tester la fiabilité des techniques de traitement du signal émergentes.

Concernant en outre les informations fournies sur la méthodologie SVM employée, 58.6% des articles n'indiquent ni une optimisation, ni les valeurs des hyperparamètres des *kernels*, ce qui nuit également à la reproductibilité. Il convient toutefois de nuancer cette observation : dans un grand nombre d'articles (*cf* par exemple [160, 233, 260] – liste non exhaustive), la méthodologie EEG employée et le protocole d'enregistrement, de la description de la tâche aux consignes données, et jusqu'au matériel dont on se sert, sont abondamment détaillés. Cette disparité dans les renseignements fournis dénote non pas une volonté de dissimuler des informations, mais plutôt un manque d'éducation aux nouvelles méthodes de *machine learning*. En effet, les neuroscientifiques ne sont que rarement formés aux techniques de traitement du signal et *a fortiori* aux méthodes les plus récentes, car ces items ne figurent pas dans les cursus universitaires de biologie. Ces lacunes dans la formation entraînent une certaine méconnaissance des paramètres critiques de la méthode expérimentale, et quant à la pertinence des informations nécessaires à la reproductibilité de leur étude.

Ce travail de thèse a consisté à se placer à la frontière entre trois domaines, les neurosciences, les mathématiques et l'informatique, en tâchant de faire le lien entre les méthodes décrites par les mathématiciens et les ingénieurs et l'interprétation biologique des résultats obtenus en *machine learning* de manière plus générale.

Notre choix s'est porté spécifiquement sur le SVM car c'est une méthode développée en mathématiques depuis 1992 [397], arrivée dans le domaine de la neuromagerie par électroencéphalographie au début des années 2010. C'est une méthode qui donc commence à être bien connue, et qui est très utilisée, comme le montre notre revue de la littérature (*cf* Chapitre 3). De plus, si le SVM est si prisé, c'est qu'il en existe déjà des fonctions dans plusieurs langages de programmation, tels que Matlab ou Python. *A contrario* des techniques plus récentes telles que le deep learning ou les réseaux de neurones, le SVM est donc moins complexe à prendre en main et surtout moins coûteux en termes de ressources de calcul nécessaires. Enfin, les collaborations précédentes des laboratoires de neurosciences et de mathématiques avaient déjà portés sur le SVM, ce qui nous donnait un certain recul par rapport à cette expérience. Nous souhaitions appliquer le SVM sur de nouvelles données cliniques afin d'examiner leur potentiel, mais aussi afin de bénéficier d'un point de comparaison lorsque nous utiliserions des méthodes de *machine learning* certes plus puissantes, mais aussi plus récentes, et donc plus difficiles à évaluer. Nous avons du toutefois effectuer un travail conséquent de programmation pour accompagner ces algorithmes à l'usage voulu en EEG.

Au cours de nos travaux, nous nous sommes également penchés sur des méthodes de visualisation des données, afin de savoir s'il était possible d'anticiper ou de confirmer les résultats du SVM.

Le t-SNE est une méthode récente, apparue dans Matlab en 2017, qui a comme particularité de pouvoir prendre en compte un grand nombre de dimensions, qui est dans notre cas la taille du vecteur de caractéristiques d'un essai, et d'en faire une projection en trois dimensions, donc visualisable. Cette façon de faire nous permet d'estimer la séparabilité des deux classes. En effet, si les deux nuages de points sont bien distincts, nous pouvons nous attendre à de bons résultats en classification, tandis que si nous constatons une forte interpénétration, les données seront plus difficilement séparables.

On peut quelque part se demander, puisqu'un simple contrôle visuel semble suffire, quel est l'intérêt d'utiliser le SVM. Le t-SNE est un algorithme très utile, car l'humain n'est pas capable de concevoir des projections dans un grand nombre de dimensions, mais qui ne permet cette projection 3D que par une certaine compression, qui peut engendrer une perte d'informations. Or le SVM, lui, peut calculer cette frontière en prenant en compte tous les espaces, ce qui lui permet de refléter l'entière réalité des données qui lui sont fournies.

Dans le chapitre 5, nous évoquons le fait que finalement, sur les EEG Rihanna, le *kernel* linéaire et l'hyperparamètre C standard suffisaient pour obtenir une performance de classification comparable à celle obtenue après optimisation. Dans ces conditions, est-il bien intéressant de toujours optimiser, au vu du temps et des ressources de calcul requises ? En réalité, la réponse à cette question va dépendre des données elles-mêmes. Dans les cas où, visuellement déjà, les données sont bien distinctes, et où nous avons donc une forte similarité intraclasse et une grande différence interclasses, il n'est pas certain qu'une optimisation puisse faire gagner beaucoup en termes de performance. Mais dans les cas où les nuages de points sont en partie superposés, l'optimisation a là un rôle à jouer pour assouplir le modèle SVM et classer correctement les données. Dans le dernier cas où nous aurions des données très peu séparables, et ce quelle que soit la méthodologie employée, l'optimisation par contre ne donnerait rien de plus probant. Mais c'est justement parce qu'il est quasiment impossible de distinguer à l'œil humain les cas limites des cas non séparables que le *machine learning* et l'optimisation des paramètres prennent toute leur place, car ils vont permettre de différencier ces deux cas. Nos tests montrent en effet que le SVM n'obtient pas une bonne performance sur les données aléatoires, et donc qu'il ne fait pas de classement sur le bruit de fond de l'EEG, lorsqu'il est bien employé. Cela est rassurant, car cela prouve que le SVM ne trouve pas des informations où il n'y en a pas. Sur le plan clinique, cela peut par exemple faire la différence entre un patient comateux dont le pronostic de réveil est mauvais et un patient dont l'état est meilleur mais indétectable avec les méthodes habituelles.

Cela soulève aussi un certain nombre de questions : quel est notre but avec le *machine learning* ? Davantage qu'un processus complètement automatique, le SVM est pour l'instant une méthode statistique qui donne un score, comme peut l'être la p-valeur pour les autres tests, mais il faut encore l'interpréter et le mettre en contexte, car seul il ne suffit pas.

Souhaite-t-on prouver qu'il y a au moins un certain nombre d'essais présentant une différence dans l'activité cérébrale ? Se contenter de détecter une différence quelconque entre les deux conditions ? Veut-on détecter un potentiel évoqué particulier ? Dans 100% des essais ? Chez 100% des volontaires sains ? Accepte-t-on de surdétecter ou de sous-détecter une des conditions par rapport à l'autre ?

Il est nécessaire de s'interroger sur ces points à chaque étude menée, afin d'être certains de se servir du *machine learning* à bon escient, car ce n'est pas une simple boîte noire avec données à l'entrée et score en sortie. Le médecin aura donc toujours un rôle crucial dans l'interprétation du résultat obtenu.

Outre les résultats présentés dans les chapitres précédents, nous avons également testé une technique de *machine learning* plus récente : l'*Echo State Network*. Il s'agit d'une méthode d'apprentissage machine bio-inspirée, basée sur l'utilisation d'un réservoir de neurones artificiels (*reservoir computing*) associés en réseaux avec différents types de connexions, dont le "poids synaptique" est modulé par le jeu de données d'entraînement

fourni au modèle [398]. Ce réservoir permet notamment l'encodage temporel et contextuel de l'information et est pour cette raison souvent assimilé à un "SVM temporel". Cette notion est très prometteuse pour l'analyse du signal EEG, de part l'excellente résolution temporelle qui fait la particularité de cette technique de neuroimagerie [399, 400]. Toutefois ces travaux en sont encore au stade préliminaire, et vont nécessiter des bases de données plus adaptées que celles dont nous disposons à l'heure actuelle.

En effet, à partir du moment où l'on travaille sur l'humain, voire le vivant d'une manière générale, un certain nombre de difficultés surgissent.

Tout d'abord, la quantité de données est assez faible. En effet, les lois de bioéthique, tout autant que les ressources allouées aux projets de recherche, ne permettent de recruter qu'un nombre limité de sujets. D'autre part, il peut être difficile d'incorporer des personnes souffrant de pathologies, en particulier mentales, pour des expériences, à cause de la rareté relative de ladite pathologie – nous avons vu notamment que les patients en état de conscience modifiés ne sont pas légion –, du stress engendré par les examens nécessités par l'étude, voire, par exemple dans le cas de la maladie d'Alzheimer, de la mise sous tutelle des patients. Ces contraintes limitent la capacité de l'échantillon de sujets à représenter la population générale, ainsi que la puissance des tests statistiques généralement pratiqués, en raison de l'existence de fortes variations inter-individuelles.

En outre, le faible ratio signal/bruit de l'EEG nous oblige à répéter un certain nombre de fois chaque tâche ou stimulation afin de pouvoir en extraire les informations essentielles. Cependant, ce processus est encore une fois coûteux et long. L'allongement de la durée de l'expérience, en particulier, représente une complication car l'humain, et surtout le patient, est fatigable. La répétition du même stimulus un grand nombre de fois est donc susceptible de provoquer une diminution de l'attention portée à celui-ci, ce qui ne fera que dégrader le ratio signal/bruit au fur et à mesure de l'enregistrement, et engendrera une différence entre les essais recueillis au début et à la fin de l'expérience [401, 402].

Une solution pourrait consister à répartir les enregistrements sur plusieurs sessions dans la même journée ou sur différents jours, comme cela a déjà été tenté dans le domaine des BCI [403]. Toutefois, ce moyen comporte lui aussi des inconvénients, dont le principal est la variation intra-individuelle : selon la température, l'humeur, l'heure de la journée, le jour *etc.*, le sujet ne présente pas la même activité cérébrale de base, ce qui conduit à un bruit de fond différent, qui est susceptible de conduire le classificateur à considérer les essais d'une même session comme semblables, par opposition à ceux d'une autre session, et ce bien que ces enregistrements soient issus de stimuli différents.

Le résultat en est que la plupart des enregistrements EEG sont à moins de 100 essais par classe, ce qui est certes raisonnable lorsque l'on a recours à des méthodes de moyennage des essais et des sujets, comme cela se fait dans les analyses de groupe traditionnelles, mais très insuffisant dès lors que l'on souhaite pratiquer une classification automatique des données.

D'autre part, si les fonctions cérébrales dites de bas niveau sont relativement aisées à explorer, les fonctions cognitives plus complexes sont plus difficiles à appréhender et nécessitent des protocoles impliquant des stimuli plus nuancés que les précédentes, comme c'est le cas dans ce travail de thèse, tant pour les études utilisant des chansons que pour celle recourant aux quatuors de mots. Ces deux tâches ont en effet deux points communs.

Le premier est que nos essais correspondent tous à des stimuli différents : dans le

cas des chansons, dans la condition “chansons connues”, nous avons cinq musiques différentes, dont les seuls points communs sont d’être chantées en langue française, et d’être connues du sujet (*cf* Partie II). À cela s’ajoute le fait que nous avons vingt périodes de silence à chaque fois, et que celles-ci sont placées à différents moments. Ainsi, deux essais “chansons connues” n’appartiennent pas nécessairement à la même chanson, et même lorsque c’est le cas, ils ne sont pas strictement identiques. La même chose a lieu avec les quatuors de mots, puisque les tétrades sont toutes différentes. Ce phénomène peut potentiellement desservir la performance de classification de par son impact sur la similarité intra-classe (*cf* Chapitre 3 section 3.3.2).

Le second point commun a trait aux consignes dispensées aux sujets lors de l’enregistrement EEG. Nous avons en effet choisi de leur demander d’effectuer la tâche de manière passive, sans se concentrer particulièrement, ceci dans le but de pouvoir plus aisément transposer le protocole aux patients, dont il est difficile de déterminer le niveau d’attention. Or il est maintenant connu que la condition attentionnelle a un impact sur l’amplitude de la réponse en potentiels évoqués [404], qui peut à son tour influencer l’accuracy de classification.

Ces limitations mettent en lumière un impératif d’évolution des protocoles de recherche en EEG afin de bénéficier de données plus nombreuses et plus facilement séparables. Nous envisageons pour cela d’utiliser dans un premier temps les enregistrements EEG réalisés quotidiennement au centre du sommeil du CHU de Besançon dans le cadre des examens cliniques polysomnographiques des patients.

Ainsi, tant les méthodes d’analyse que les tâches effectuées par les sujets pour les examens EEG seront amenées vers de nouvelles normes, marquant une étape supplémentaire dans l’évolution de l’EEG (*cf* Chapitre 2.1). Il est à se demander également si l’appareil EEG lui-même ne tirerait pas avantage de certaines améliorations, notamment dans l’élaboration de nouvelles électrodes moins vulnérables au bruit environnant, qui permettraient d’augmenter le ratio signal sur bruit à la racine. Des tentatives ont déjà été effectuées dans ce sens [405], mais sont encore récentes et doivent être soumises à davantage de tests avant de pouvoir être commercialisées.

7.2/ QUELQUES CONSIDÉRATIONS SUR L'IA EN MÉDECINE

7.2.1/ VERS UNE ANALYSE DES DONNÉES PLUS PERFORMANTE ET REPRODUCIBLE

Malgré ces limitations, les systèmes basés sur l’apprentissage machine ont fait leurs preuves dans divers essais de recherche, et leur impact potentiel sur les soins de santé est profond, en particulier dans des domaines tels que le diagnostic et l’imagerie médicale.

Leur principal avantage est de faciliter des tâches certes intellectuelles mais répétitives, telles que l’étiquetage des différentes phases du sommeil en EEG, qui nécessitent une réelle expertise de la part du clinicien, mais qui sont difficilement reproductibles en raison des variations inter-individuelles existant d’un expert à un autre, mais aussi intra-individuelles, un même expert n’étant pas assuré, face aux mêmes données, de les évaluer chaque fois à l’identique. Cette problématique est connue depuis longtemps [406, 407], mais aujourd’hui l’apprentissage machine obtient des résultats tout à fait honorables,

parfois même en concurrence avec ces experts [408, 409]. De plus, et c'est loin d'être négligeable, plus les modèles sont formés à reconnaître des cas spécifiques et précis, meilleure est alors leur performance.

D'autre part, l'utilisation de ces nouvelles méthodes d'analyse permet de dégager un temps précieux pour le clinicien, et est moins coûteuse à terme que la formation et la mobilisation d'un expert humain.

7.2.2/ DES DISCOURS ALARMISTES DANS LES MÉDIAS, MAIS FINALEMENT PEU FONDÉS

L'intelligence artificielle et son impact sur l'avenir des soins de santé font l'objet d'un grand battage médiatique et suscitent beaucoup de craintes dans certaines professions médicales, notamment que l'IA, et l'automatisation en général, ne leur prennent leur emploi. Dans le domaine de l'imagerie médicale en particulier, de nombreux radiologues se sont mis à paniquer. Dans sa présentation à la GPU Tech Conference à San José (USA) en mai 2017, Curtis Langlotz², professeur de radiologie et d'informatique biomédicale à l'université de Stanford, mentionne comment il a reçu un courriel de l'un de ses étudiants, disant qu'il pensait se lancer dans la radiologie mais ne sait pas si c'est encore une profession viable.

L'investisseur Vinod Khosla, de la Silicon Valley, est allé jusqu'à déclarer que les machines remplaceront 80% des médecins à l'avenir, dans un secteur de la santé dirigé par des entrepreneurs, et non par des professionnels de la santé³.

Toutefois, les investisseurs tels que Vinod Khosla ne comprennent pas entièrement le secteur des soins de santé, et selon Frey *et al.* (2013) [410], il semble que sa vision ne se concrétisera pas, puisque si les transcritteurs médicaux, les techniciens des dossiers médicaux et de l'information sanitaire et les secrétaires médicales sont les emplois les plus susceptibles d'être informatisés à l'avenir, les médecins et les chirurgiens n'ont en revanche que 0,42% de chances que leur profession soit automatisée.

Ainsi, cette vision des choses ne prend pas forcément en compte la complexité du diagnostic et de la relation humaine s'établissant entre le soignant et le soigné.

7.2.3/ POURQUOI LE MÉDECIN N'EST TOUT SIMPLEMENT PAS REMPLAÇABLE

Créativité et capacité d'interprétation

Bien que les données, les mesures et les analyses quantitatives soient une partie cruciale du travail d'un médecin, l'établissement d'un diagnostic et le traitement d'un patient sont des processus non linéaires qui requièrent une créativité et des capacités de résolution de problèmes que les algorithmes et les robots ne possèdent pas. Jusqu'à l'émergence de solutions numériques complexes, les médecins traduisaient les données provenant de simples appareils médicaux en décisions médicales. À l'avenir, la tâche sera la même, ils n'utiliseront que des technologies plus perfectionnées.

2. GTC conference S7508 (2017), par Curtis Langlotz, <https://blogs.nvidia.com/blog/tag/gtc-2017/>

3. *Do we need doctors or algorithms?* (2012), par Vinod Khosla, <https://techcrunch.com/2012/01/10/doctors-or-algorithms/>

Nécessité d'experts formés aux nouveaux outils

Des solutions de santé numériques de plus en plus sophistiquées nécessiteront la compétence de professionnels médicaux qualifiés. Par exemple, le logiciel d'IBM Watson destiné aux oncologues offre aux cliniciens des options de traitement fondées sur des preuves. Néanmoins, seuls les médecins et leurs patients peuvent choisir le traitement, et seuls les médecins peuvent évaluer si l'algorithme a fourni des suggestions potentiellement utiles. Aucun algorithme ne peut interpréter clairement des défis complexes impliquant la psyché. S'ils fournissent les données, l'interprétation restera toujours un territoire humain.

7.2.4/ MAIS UN OPTIMISME BIEN PRÉSENT CHEZ LE PERSONNEL MÉDICAL

Afin de savoir ce que les médecins pensent de la mise en œuvre potentielle de l'IA en médecine, Sarwar *et al.* (2019) [411] ont mené une enquête auprès de 487 pathologistes dans 54 pays.

Dans l'ensemble, les médecins interrogés se sont montrés optimistes quant à l'intégration de l'IA et n'ont guère eu peur de perdre leur emploi. Au contraire, 42,4% des médecins interrogés ont estimé que l'introduction de l'IA en médecine créerait de nouveaux postes et augmenterait l'emploi.

La plupart des praticiens interrogés estiment que les outils d'IA pourraient accroître l'efficacité du diagnostic. Bien qu'ils soient très optimistes quant à la capacité de la technologie, 48,3% estiment que le diagnostic des patients devrait toujours être effectué principalement par le médecin et 25,3% pensent que la tâche pourrait être partagée de manière égale entre l'homme et l'IA. Seule une minorité de répondants estime que l'IA devrait jouer un rôle dominant dans le processus décisionnel. De nombreuses réponses indiquant qu'elles reconnaissent le risque d'erreur mécanique, le consensus semble être que l'IA devrait être appliquée comme un outil supplémentaire dans le processus de diagnostic.

7.2.5/ VERS UNE TRANSFORMATION DE L'EXERCICE DE LA MÉDECINE

Autrement dit, il ne semble pas se profiler de confrontation entre la technologie et l'humain, mais plutôt une collaboration. Une étude visant à identifier le cancer du sein métastatique grâce à du deep learning montre que lorsque les prédictions du système d'apprentissage profond ont été combinées aux diagnostics du médecin humain, la performance de classification des images, ainsi que le score de localisation de la tumeur, ont augmenté de manière significative, tandis que le taux d'erreur humaine diminuait de 85% [412].

Il ne fait aucun doute que les algorithmes joueront un rôle croissant dans la routine quotidienne des soins de santé pour diagnostiquer les cas les plus simples et prendre en charge des tâches répétitives. Cela devrait permettre de dégager davantage de temps pour le praticien afin de tirer parti de son expérience clinique et émotionnelle pour établir des relations approfondies avec ses patients. En effet, un diagnostic, même s'il est posé avec une rapidité stupéfiante, ne suffit pas. Nous commençons à peine à comprendre dans quelle mesure la guérison du patient peut être influencée positivement par le lien entre le patient et le médecin, mais un nombre croissant de recherches suggère que

cette possibilité pourrait être importante [413]. La grande majorité des gens continue aujourd'hui d'accorder une grande importance à la possibilité de se faire soigner par un médecin qu'ils connaissent et avec lequel ils ont une relation de confiance.

C'est dans cet état de fait que réside une importante opportunité : utiliser la technologie pour rendre le temps passé avec les patients beaucoup plus significatif et cliniquement efficace que jamais auparavant.

CONCLUSION

Les objectifs de ce travail de thèse étaient de transposer les méthodes d'apprentissage machine au traitement du signal EEG, et plus particulièrement le *Support Vector Machine* ; et d'évaluer dans quelles mesures ces nouvelles méthodes permettent un niveau individuel d'analyse des données par rapport aux méthodes traditionnelles pour le diagnostic différentiel.

Dans le chapitre 3, nous avons passé en revue 145 articles appliquant le SVM à l'EEG, couvrant différents domaines tels que les BCI, le diagnostic, le monitoring cognitif et émotionnel *etc...* afin d'en dégager les tendances et d'évaluer les limites des méthodes déjà employées pour formuler des recommandations. Notre analyse révèle que plus de 70% des études utilisent des données privées, recueillies spécifiquement pour l'étude en question et donc non accessibles au public. Si les processus de pré-traitement des données sont relativement transparents, fournis et discutés dans les articles, nous avons constaté en revanche que les paramètres liés à la structure de classification, i.e. la méthode de validation croisée, les valeurs retenues pour les hyperparamètres et le choix d'optimiser ceux-ci ou non ne sont que rarement fournis (60 études sur 145 seulement). Mais surtout, nous avons remarqué que les études souffrent souvent d'une mauvaise reproductibilité : une majorité d'articles serait difficile, voire impossible à reproduire étant donné l'indisponibilité de leurs données et de leur code.

Le projet Rihanna, présenté en chapitre 4, avait pour objectif de discriminer les patients atteints de la maladie d'Alzheimer des patients souffrants de la dépression. Dans ce but, les capacités mnésiques de sujets sains ont été mesurées lors de l'écoute de 5 chansons connues et de 5 chansons inconnues dans lesquelles avaient été insérés des silences, l'étude de Gabriel *et al.* (2016) [292], menée au sein de notre laboratoire, ayant montré qu'il était possible d'évaluer la familiarité musicale à l'aide de l'imagerie musicale. S'agissant d'un protocole novateur, une des plus importantes étapes nous semblait être d'explorer les biais influençant possiblement l'apprentissage de nouvelles chansons. Dans ce contexte, l'objectif de cette étude était d'étudier si les neuromarqueurs de l'apprentissage sont affectés ou non par l'appréciation musicale. Nous avons ainsi exploré les mécanismes neuronaux de l'appréciation musicale chez vingt volontaires sains. Nous avons mis en évidence une réponse spécifique à court terme, impliquant une modulation de l'amplitude des potentiels évoqués MMN et P3a par l'appréciation musicale. Sur l'ensemble de la période de silence, des activités oscillatoires spécifiques sont apparues, se traduisant par des synchronisations/désynchronisations de l'activité thêta et alpha au sein de structures cérébrales particulières telles que les *gyri* frontal inférieur et temporal supérieur.

Dans le chapitre 5, nous avons appliqué la classification SVM aux données du projet Rihanna afin de discriminer automatiquement l'activité cérébrale induit par l'écoute de chansons connues et inconnues. Nous avons proposé un pipeline consistant en une étape préliminaire de prétraitement des données à l'aide de la méthode xDAWN, puis en l'exploration de la meilleure façon d'extraire les informations pertinentes des enregistrements EEG tant dans le domaine temporel que fréquentiel. Enfin, notre étude souligne la nécessité d'optimiser les hyperparamètres du SVM pour chaque sujet d'intérêt.

Dans le chapitre 6, nous avons pour objectif d'examiner la possibilité de classer automatiquement et de façon individualisée les potentiels évoqués liés au traitement présémantique. Notre hypothèse était que l'optimisation individualisée de l'extraction des caractéristiques par le biais d'un algorithme génétique améliorerait nettement la précision de la classification. Nous avons confirmé la présence au niveau de l'analyse de groupe des réponses attentionnelles et cognitives attendues, dont la N400. Nos analyses ont révélé que l'amplitude et la durée de ce potentiel évoqué sont bien sensibles à l'incongruence au contexte du quatrième mot. L'ensemble des résultats obtenus lors de la classification pointe vers la confirmation d'une forte variabilité inter-individuelle, comme en témoignent plus particulièrement les fluctuations de période d'intérêt, de taille de celle-ci, ainsi que de longueur de la fenêtre glissante et du chevauchement d'un sujet à l'autre.

Le chemin est encore long avant que nous soyons capables de proposer des méthodes efficaces et surtout fiables, utilisables en routine clinique dans le cadre du diagnostic différentiel des pathologies neuropsychiatriques.

Il apparaît indispensable de définir des protocoles d'investigation adaptés aux contraintes spécifiques du *machine learning* : un bien plus grand nombre d'essais, et des stimuli dont les permettant l'induction d'activités cérébrales plus aisément détectables.

Mais surtout, nous insistons sur la nécessité pour la communauté scientifique d'investir ses efforts non seulement dans de nouvelles techniques de traitement du signal, mais également dans la compréhension et la formation des neuroscientifiques à l'utilisation de celles-ci.

Dans ce contexte, la mise à disposition des données enregistrées et des programmes d'analyse et de classification peut sans aucun doute participer à une meilleure reproductibilité des études menées en *machine learning*, en offrant la possibilité à d'autres équipes de recherche de tester et de discuter des performances de classification obtenues sur des bases communes.

Une fois ces bases posées, nous pourrons alors travailler à établir un consensus autour des protocoles et des techniques de traitement du signal les mieux adaptées au diagnostic de telle ou telle pathologie.

Enfin, les expérimentateurs appliquant du *machine learning* à des données de santé doivent s'engager à davantage vulgariser leurs travaux afin de les démystifier. Les médecins tout autant que les patients doivent exiger d'être tenus au courant des bénéfiques potentiels comme des limitations de cette technologie dans le diagnostic et les soins, afin de réduire l'anxiété liée à la dramatisation médiatique de l'intelligence artificielle.

Une fois ces considérations ancrées, les avancées technologiques de l'électroencéphalographie, couplées à ces méthodes de traitement et de classification automatique du signal ouvrent pour le monde médical de tout nouveaux horizons, qu'il ne nous reste plus qu'à explorer.

TABLE DES FIGURES

1.1	En bleu, l'amygdale fonctionne en lien avec l'hippocampe et le cortex entorhinal lors de la mise en mémoire d'événements à connotation émotionnelle. Elle est impliquée tant dans la maladie d'Alzheimer que dans la dépression. En orange, l'hippocampe est la première région lésée lors de la neurodégénérescence induite par la maladie d'Alzheimer. Image générée à partir d'Anatomography, site web administré par Life Science Databases. https://lifesciencedb.jp/bp3d/#	10
1.2	Les différentes étapes de récupération. Un patient en état végétatif (UWS) va généralement évoluer vers un état de conscience minimale (MCS), puis éventuellement émerger de cet état vers une récupération complète ou avec séquelles neurologiques ou psychiques. Tiré de Bruno <i>et al.</i> (2011). [46]	15
2.1	Vue de profil d'un casque EEG haute résolution à 256 électrodes (Hydro-Cel Geodesic Sensor Net, EGI).	23
2.2	Illustration de la génération de très petits champs électriques par des courants synaptiques dans des cellules pyramidales. L'électrode EEG mesure le signal à travers les épaisses couches de tissus. Ce n'est que si des milliers de cellules présentent simultanément leurs petites tensions que les signaux deviennent suffisamment importants pour être visibles à la surface du cuir chevelu. Traduit de Siuly, Li, et Zhang (2016) [75].	25
2.3	Illustration de la résolution spatio-temporelle des technologies actuelles de monitoring du cerveau. Adapté et traduit de Muthuswamy, Sridharan et Okandan (2016). [95]	28
3.1	Représentation à deux dimensions d'une frontière de séparation construite par le SVM. A. Essai bien classé et pas dans la zone de la marge. Il ne représente pas un vecteur support. B. Essai bien classé et aux frontières de la zone de la marge. Il représente un vecteur support. C. Essai bien classé et dans la zone de la marge. Il ne représente pas un vecteur support. D. Essai du mauvais côté de la frontière, il est mal classé et ne représente pas un vecteur support.	34
3.2	Illustration de l'astuce du noyau : les données sont projetées dans un espace de plus grande dimension, où les données peuvent être séparées par un hyperplan. Cette frontière linéaire dans l'espace de grande dimension est équivalente à une frontière non linéaire dans l'espace d'origine.	35

3.3	Représentation à deux dimensions du type de frontière de séparation possible par l'utilisation de kernels. a) <i>Kernel</i> linéaire. b) <i>Kernel</i> polynomial. c) <i>Kernel</i> RBF.	35
3.4	Artéfacts les plus couramment rencontrés dans l'analyse du signal EEG. a) Artefact électro-oculographique causé par l'excitation des muscles du globe oculaire (lié au clignement des yeux, par exemple). Grande amplitude, lente, onde positive proéminente dans les électrodes frontales. b) Artefact d'électrode causé par un mauvais contact (et donc une plus grande impédance) entre l'électrode P3 et la peau. c) Artefact de déglutition. d) Artefact de l'électrode de référence commune causé par un mauvais contact entre l'électrode de référence et la peau. Onde énorme similaire dans tous les canaux. Tiré de : https://en.wikipedia.org/wiki/Electroencephalography .	38
3.5	Matrice de confusion permettant de calculer les différentes métriques de performance. Précision = (Vrais positifs + Vrais négatifs) / Nombre total d'essais. Spécificité = Vrais négatifs / (Vrais négatifs + Faux positifs). Sensibilité = Vrais positifs / (Vrais positifs + Faux négatifs).	40
3.6	Architectures de validation croisée les plus utilisées. a) Validation croisée <i>k-fold</i> : le jeu d'entraînement est divisé en sous-ensemble, dont chacun est utilisé à son tour pour valider le modèle produit par l'entraînement du modèle SVM sur tous les autres sous-ensembles. b) Validation croisée <i>Leave-One-Out</i> , ou <i>Leave-One-Subject-Out</i> : cette fois le sous-ensemble de validation du modèle ne comporte qu'un seul essai ou sujet.	40
3.7	Nombre d'articles passés au crible des critères d'inclusion et d'exclusion. Cent quarante-cinq études ont été incluses dans notre analyse.	44
3.8	Pays d'origine de l'affiliation (laboratoire, institut ou université) du premier auteur. Carte générée à partir de https://mapchart.net/world.html	45
3.9	Répartition des articles par année et domaine de publication. a) Nombre d'articles publiés par année, de 2004 à mai 2020. b) Nombre d'articles publiés par année et par domaine. Pour faciliter la lecture de l'histogramme, les catégories comprenant moins de 5 publications n'ont pas été incluses dans la figure.	45
3.10	Domaine principal de chaque étude. Le nombre d'articles de chaque catégorie est indiqué entre parenthèses. Les études traitant de plus d'un sujet ont été classées selon leur thème principal.	47
3.11	Différentes approches sont utilisées dans la gestion des sujets : classification intra-sujet, inter-sujets ou single-trial (ST). a) Nombre de publications par type d'approche. b) Nombre d'essais utilisés par sujet dans chacune des trois approches. c) Nombre de sujets utilisés dans chacune des trois approches.	48
3.12	Caractéristiques du matériel et fréquence d'échantillonnage. a) Nom et fabricant du matériel EEG utilisé. b) Nombre d'électrodes conservées pour l'analyse. c) Fréquence d'échantillonnage utilisée pour l'analyse (en Hertz).	49

3.13	Choix effectués concernant le pré-traitement EEG. a) Pourcentage d'articles ayant appliqué un pré-traitement (filtrage, interpolation d'électrodes etc...) ou non, ou n'ayant rien spécifié (NS). b) Pourcentage d'articles ayant utilisé une méthode automatique ou humaine, ou n'ayant rien spécifié concernant la conservation, l'élimination ou la correction des artéfacts (NS). c) Pourcentage d'articles ayant utilisé le signal EEG brut, extrait des caractéristiques du domaine fréquentiel, une combinaison des deux, d'autres caractéristiques, ou n'ayant rien spécifié (NS). d) Pourcentage d'articles ayant appliqué une normalisation des données lors de la classification (z-score ou autre), ou n'ayant pas spécifié de normalisation quelconque (NS).	51
3.14	Choix effectués concernant les hyperparamètres du SVM. a) Pourcentage d'articles ayant ou non utilisé une méthode d'optimisation des hyperparamètres du SVM. b) Pourcentage d'articles ayant ou non précisé les valeurs retenues pour les hyperparamètres du SVM. c) Pourcentage d'articles ayant ou non proposé l'un ou l'autre. d) Nombre d'articles ayant choisi tel ou tel <i>kernel</i> (NS : non spécifié).	52
3.15	Métriques de performance et méthodologie de validation croisée retenues. a) Métriques de performance utilisées dans les articles analysés. Un même article peut avoir employé plusieurs d'entre elles. b) Approche retenue pour la validation croisée (NS : non spécifiée).	52
3.16	Reproductibilité des études examinées. a) Disponibilité du code utilisé pour le pré-traitement et/ou la classification. b) Disponibilité des données EEG utilisées dans les articles. c) Estimation du niveau de reproductibilité des études (Facile : code et données disponibles, Moyen : code ou données disponibles sur demande, Difficile : données disponibles mais pas le code, Impossible : ni les données ni le code ne sont disponibles).	53
3.17	Déroulement de la procédure expérimentale.	60
3.18	Schéma représentant le déroulement de l'expérience EEG. Le début et la fin de chaque silence est repéré par un marqueur différent afin de retrouver ces périodes sur le tracé EEG. 20 silences d'une durée de 2 secondes ont été introduits, avec un délai variant de 2.5 à 3.5 s entre les périodes de silence. Les 10 chansons sont présentées à l'écoute de manière aléatoire, deux fois de suite. Des pauses sont proposées au sujet entre les sessions.	61
4.1	Déroulement de la procédure expérimentale.	65
4.2	The nine regions of interest. A: Anterior, M: Median, P: Posterior; L: Left, C: Central, R: Right.	69
4.3	Time frequency analysis performed over the nine ROIs during silence inserted into liked songs. Neural synchronizations (yellow) and desynchronizations (blue) are represented at different frequency bands. Areas of significant differences ($p < 0.01$) between liked and disliked songs are surrounded in red.	71

4.4	Brain areas showing significant differences in activation between liked and disliked songs. Left: source of difference in theta activity in the first time period. Middle: source of difference in theta activity in the second time period ($p < 0.001$). Right: source of difference in alpha activity in the last time period ($p < 0.001$). Areas more strongly activated for liked songs are represented in red, and in blue for disliked songs.	72
4.5	ERPs for liked and disliked songs in the three anterior ROIs. A significant effect of gradient was obtained, with neural differences in musical preference on anterior ROIs for both TW2 and TW3.	76
5.1	Regions of Interest (ROIs) used for classification, based on Gabriel <i>et al.</i> (2016). [292]	82
5.2	Preprocessing and classification pipeline proposed for each subject individually.	85
5.3	Proportion of subjects for whom a particular features extraction method is best regarding classification accuracy, for both linear and RBF kernels. a) Methods using all of the 257 channels. b) Methods using ROIs. c) Best over 257 channels and ROIs.	88
5.4	Linear kernel: evolution of training, validation and testing accuracies as a function of the hyperparameter C. a) Profile A. b) Profile B. c) Number of subjects who belong to profile A or B for each features extraction method.	91
5.5	RBF kernel: evolution of training, validation and testing accuracies as a function of the hyperparameters C and sigma. a) Training accuracy, profile A. b) Training accuracy, profile B. c) Validation accuracy, profile A. d) Validation accuracy, profile B. e) Testing accuracy, profile A. f) Testing accuracy, profile B. g) Number of subjects who belong to profile A or B for each features extraction method.	93
5.6	Linear kernel: density of chosen hyperparameter C across 50 permutations, for roiMean method. a) Number of times a hyperparameter was chosen over all 50 permutations for all subjects. b) Number of times a hyperparameter was chosen over all 50 permutations for each subject individually.	94
5.7	Linear kernel: density of chosen hyperparameter C across 50 permutations, for roi4bFFT method. a) Number of times a hyperparameter was chosen over all 50 permutations for all subjects. b) Number of times a hyperparameter was chosen over all 50 permutations for each subject individually.	95
5.8	Linear kernel: density of chosen hyperparameter C across 50 permutations, for all4bFFT method. a) Number of times a C value was chosen over all 50 permutations for all subjects. b) Number of times a C value was chosen over all 50 permutations for each subject individually.	96
5.9	RBF kernel: density of chosen pair of hyperparameters C and sigma across 50 permutations, for roiMean method. a) Number of times a sigma value was chosen over all 50 permutations for each subject individually. b) Number of times a pair (C, sigma) was chosen over all 50 permutations for all subjects. c) Number of times a C value was chosen over all 50 permutations for each subject individually.	97

5.10 RBF kernel: density of chosen hyperparameter C across 50 permutations, for roi4bFFT method. a) Number of times a sigma value was chosen over all 50 permutations for each subject individually. b) Number of times a pair (C, sigma) was chosen over all 50 permutations for all subjects. c) Number of times a C value was chosen over all 50 permutations for each subject individually.	97
5.11 RBF kernel: density of chosen hyperparameter C across 50 permutations, for all4bFFT method. a) Number of times a sigma value was chosen over all 50 permutations for each subject individually. b) Number of times a pair (C, sigma) was chosen over all 50 permutations for all subjects. c) Number of times a C value was chosen over all 50 permutations for each subject individually.	98
6.1 Exemple de deux quatuors de mots. a) Le quatrième mot est congruent au contexte sémantique fixé par les trois premiers, ici les arts musicaux. b) Le quatrième mot est incongruent au contexte sémantique : c'est une couleur.	110
6.2 Extraction des caractéristiques fréquentielles du signal EEG par FFT à fenêtre glissante. La fenêtre se déplace tout au long de la période d'intérêt sélectionnée dans l'essai EEG, avec un certain chevauchement entre deux fenêtres. Cette méthode permet d'avoir une représentation temps-fréquence et non uniquement fréquentielle du signal.	112
6.3 Pipeline incluant l'extraction des caractéristiques et la classification des essais au sein de l'algorithme d'optimisation génétique pour chacun des sujets.	115
6.4 Grandes moyennes de l'ensemble des sujets pour les conditions congruente et incongruente et cartographies du signal des 63 électrodes. Six topographies ont été établies après la réalisation de la segmentation, montrant l'émergence de différentes périodes de stabilité, correspondant à des mécanismes cognitifs particuliers à un moment donné.	116
6.5 Cartographies du signal EEG dans le domaine fréquentiel, obtenues sur les grandes moyennes de l'ensemble des sujets, pour chacun des bandes de fréquences considérées et pour a) la condition congruente et b) la condition incongruente.	117
6.6 Analyse temps-fréquence sur le GFP de chacune des deux conditions (a) mots congruents et b) mots incongruents). Une forte synchronisation (rouge) suivie d'une désynchronisation (jaune et bleu) est clairement observée pour chaque catégorie. Une différence dans la synchronisation entre les deux conditions est observée en c).	118
6.7 foo bar	119
6.8 Position de la période d'intérêt optimale retenue pour chaque sujet par l'algorithme génétique. La courbe orange représente le pourcentage de sujets pour lesquels chaque milliseconde fait partie de la période optimale.	119

6.9 Paramètres optimaux obtenus par algorithme génétique. a) Longueur en temps de la période d'intérêt, de la fenêtre glissante et du chevauchement optimaux pour chaque sujet. b) Nombre de sujets par longueur optimale de la période d'intérêt. c) Nombre de sujets par ratio longueur optimale de la fenêtre glissante sur longueur optimale de la période d'intérêt. d) Nombre de sujets par longueur optimale du chevauchement sur longueur optimale de la fenêtre glissante.	120
6.10 Optimisation sur données générées aléatoirement. Précision de validation obtenue par optimisation par algorithme génétique en fonction du nombre d'essais par classe.	121

LISTE DES TABLES

1.1	Critères diagnostiques de la démence de type Alzheimer. Le début de la démence Alzheimer implique un déclin cognitif progressif et continu. Son diagnostic est un diagnostic d'exclusion. Tableau tiré du Manuel diagnostique et statistique des troubles mentaux IV (texte révisé) [2].	8
1.2	Critères diagnostiques de l'épisode dépressif majeur. La caractéristique principale de l'épisode dépressif est une humeur dépressive ou une perte d'intérêt ou de plaisir pour presque toutes les activités pendant au moins deux semaines. Tableau tiré du Manuel diagnostique et statistique des troubles mentaux IV (texte révisé) [2].	12
1.3	Difficultés dans le diagnostic différentiel maladie d'Alzheimer et dépression chez le sujet âgé. Des symptômes de la maladie d'Alzheimer (en orange) peuvent apparaître dans le tableau clinique du patient dépressif (en bleu) et réciproquement.	13
1.4	Taxonomie britannique des pathologies persistantes majeures de la conscience. Traduit de Zeman (2008). [49]	18
3.1	Critères d'inclusion et d'exclusion des articles passés en revue.	42
3.2	Catégories et description des informations extraites de chacun des articles sélectionnés : origine de l'article, résumé de l'étude, données utilisées, pré-traitement EEG, méthodologie SVM, résultats obtenus et reproductibilité.	43
4.1	Musical appreciation, familiarity, and emotions for all 5 songs, as well as for the most liked and disliked songs (mean followed by standard deviation). P value shows the difference of scores between liked and disliked songs.	70
5.1	Summary of feature extraction methods: method, name abbreviation, feature type and features vector size.	83
5.2	Best testing accuracy based on validation for each subject and kernel. The best accuracy is determined across all ROIs and features extraction methods examined. Third and sixth columns: in blue colour, significance of the pValue obtained by permutation test. Logical indicator, if the accuracy is higher than 55.6%, flag = 1, otherwise flag = 0, according to the Müller-Putz (2008) accuracy significance test [353].	86
5.3	Best testing accuracy and kernel based on validation for each subject. Computation times ratio between kernels is given by RBF kernel CPU time divided by linear kernel CPU time.	87

5.4	Linear kernel: a 3-factor analysis of variance (ANOVA) was performed on each subject for each ROI: Frontal position (3 levels: front, median and back), Sagittal position (3 levels: left, central and right), features extraction method (6 levels: roiGFP, roiMean, roi4bFFT, roiMean4bFFT, roiWhFFT and roiMeanWhFFT). The post-hoc comparison of inter-category interactions was performed with a Bonferroni correction. The significance of the results was set at $p < 0.05$	89
5.5	RBF kernel: a 3-factor analysis of variance (ANOVA) was performed on each subject for each ROI: Frontal position (3 levels: front, median and back), Sagittal position (3 levels: left, central and right), features extraction method (6 levels: roiGFP, roiMean, roi4bFFT, roiMean4bFFT, roiWhFFT and roiMeanWhFFT). The post-hoc comparison of inter-category interactions was performed with a Bonferroni correction. The significance of the results was set at $p < 0.05$	90
5.6	Linear kernel, for each subject: best testing accuracy based on validation, testing accuracy based on standard $C = 1$ and best testing accuracy based on test.	99
5.7	RBF kernel, for each subject: best testing accuracy based on validation, testing accuracy based on standard $C = 1$ and $\sigma = 1$, and best testing accuracy based on test.	100
5.8	Results of classification pipeline on randomly generated data, for both linear and RBF kernels, for each features extraction method. Size of features vectors were preserved for each features extraction method for fair comparison. Best testing accuracy based on validation, testing accuracy based on standard hyperparameters and best testing accuracy based on test. . .	101
6.1	Corrélations de Spearman (ρ) entre les différents paramètres optimisés et la précision de test obtenue sur les essais préservés.	121

BIBLIOGRAPHIE

- [1] LOBO, A., LAUNER, L. J., FRATIGLIONI, L., ANDERSEN, K., DI CARLO, A., BRETELLER, M. M., COPELAND, J. R., DARTIGUES, J. F., JAGGER, C., MARTINEZ-LAGE, J., SOININEN, H., AND HOFMAN, A. **Prevalence of dementia and major subtypes in Europe: A collaborative study of population-based cohorts. Neurologic Diseases in the Elderly Research Group.** *Neurology* 54 (2000), S4–S9.
- [2] ASSOCIATION, A. P. **DSM-IV-TR Manuel Diagnostique et Statistique Des Troubles Mentaux: Texte Révisé.** Elsevier Masson, 2004.
- [3] EICHENBAUM, H., AND COHEN, N. J. **Memory, Amnesia, and the Hippocampal System.** MIT Press, 1993.
- [4] SQUIRE, L. R. **The neuropsychology of memory.** In *The Biology of Learning.* Springer Berlin Heidelberg, 2002, pp. 667–686.
- [5] SARAZIN, M., CHAUVIRÉ, V., GERARDIN, E., COLLIOT, O., KINKINGNÉHUN, S., DE SOUZA, L. C., HUGONOT-DIENER, L., GARNERO, L., LEHÉRICY, S., CHUPIN, M., AND DUBOIS, B. **The amnesic syndrome of hippocampal type in Alzheimer's disease: an MRI study.** *Journal of Alzheimer's disease : JAD* 22, 1 (2010), 285–294.
- [6] BONTEMPI, B., LAURENT-DEMIR, C., DESTRADE, C., AND JAFFARD, R. **Time-dependent reorganization of brain circuitry underlying long-term memory storage.** *Nature* 400, 6745 (1999), 671–675.
- [7] DUDAI, Y. **The restless engram: consolidations never end.** *Annual Review of Neuroscience* 35 (2012), 227–47.
- [8] MOSCOVITCH, M., AND NADEL, L. **Consolidation and the hippocampal complex revisited: in defense of the multiple-trace model.** *Current Opinion in Neurobiology* 8, 2 (1998), 297–300.
- [9] WITTENBERG, G. M., AND TSIEN, J. Z. **An emerging molecular and cellular framework for memory processing by the hippocampus.** *Trends in neurosciences* 25, 10 (2002), 501–505.
- [10] LUI, J. H., HANSEN, D. V., AND KRIEGSTEIN, A. R. **Development and evolution of the human neocortex.** *Cell* 146, 1 (2011), 18–36.
- [11] BRAAK, H., TREDICI, K. D., BOHL, J., BRATZKE, H., AND BRAAK, E. **Pathological Changes in the Parahippocampal Region in Select Non-Alzheimer's Dementias.** *Annals of the New York Academy of Sciences* 911, 1 (2006), 221–239.
- [12] PHELPS, E. A. **Human emotion and memory: interactions of the amygdala and hippocampal complex.** *Current Opinion in Neurobiology* 14, 2 (2004), 198–202.

- [13] COSTAFREDA, S. G., BRAMMER, M. J., DAVID, A. S., AND FU, C. H. Y. **Predictors of amygdala activation during the processing of emotional stimuli: a meta-analysis of 385 PET and fMRI studies.** *Brain Research Reviews* 58, 1 (2008), 57–70.
- [14] RICHARDSON, R. T., AND DELONG, M. R. **Electrophysiological studies of the functions of the nucleus basalis in primates.** *Advances in experimental medicine and biology* 295 (1991), 233–52.
- [15] LIU, R., CRAWFORD, J., CALLAHAN, P. M., TERRY, A. V., CONSTANTINIDIS, C., AND BLAKE, D. T. **Intermittent stimulation in the nucleus basalis of meynert improves sustained attention in rhesus monkeys.** *Neuropharmacology* 137 (2018), 202–210.
- [16] RIDLEY, R. M., BAKER, H. F., LEOW-DYKE, A., AND CUMMINGS, R. M. **Further analysis of the effects of immunotoxic lesions of the basal nucleus of Meynert reveals substantial impairment on visual discrimination learning in monkeys.** *Brain Research Bulletin* 65, 5 (2005), 433–442.
- [17] LIU, A. K. L., CHANG, R. C.-C., PEARCE, R. K. B., AND GENTLEMAN, S. M. **Nucleus basalis of Meynert revisited: anatomy, history and differential involvement in Alzheimer's and Parkinson's disease.** *Acta Neuropathologica* 129, 4 (2015), 527–540.
- [18] NESTLER, E. J., HYMAN, S. E., AND MALENKA, R. C. **Molecular neuropharmacology: a foundation for clinical neuroscience.** New York : McGraw-Hill Medical, 2009, ch. Chapter 6: Widely Projecting Systems: Monoamines, Acetylcholine, and Orexin, p. 157.
- [19] BONDAREFF, W., MOUNTJOY, C. Q., AND ROTH, M. **Loss of neurons of origin of the adrenergic projection to cerebral cortex (nucleus locus ceruleus) in senile dementia.** *Neurology* 32, 2 (1982), 164–168.
- [20] HENEKA, M. T., RAMANATHAN, M., JACOBS, A. H., DUMITRESCU-OZIMEK, L., BILKEI-GORZO, A., DEBEIR, T., SASTRE, M., GALLDIKS, N., ZIMMER, A., HOEHN, M., HEISS, W.-D., KLOCKGETHER, T., AND STAUFENBIEL, M. **Locus ceruleus degeneration promotes Alzheimer pathogenesis in amyloid precursor protein 23 transgenic mice.** *The Journal of neuroscience : the official journal of the Society for Neuroscience* 26, 5 (2006), 1343–1354.
- [21] HENEKA, M. T., NADRIGNY, F., REGEN, T., MARTINEZ-HERNANDEZ, A., DUMITRESCU-OZIMEK, L., TERWEL, D., JARDANHAZI-KURUTZ, D., WALTER, J., KIRCHHOFF, F., HANISCH, U.-K., AND KUMMER, M. P. **Locus ceruleus controls Alzheimer's disease pathology by modulating microglial functions through norepinephrine.** *Proceedings of the National Academy of Sciences* 107, 13 (2010), 6058–6063.
- [22] PEYRON, C., PETIT, J. M., RAMPON, C., JOUVET, M., AND LUPPI, P. H. **Fore-brain afferents to the rat dorsal raphe nucleus demonstrated by retrograde and anterograde tracing methods.** *Neuroscience* 82, 2 (1998), 443–468.
- [23] COPELAND, J. R., BEEKMAN, A. T., DEWEY, M. E., HOOIJER, C., JORDAN, A., LAWLOR, B. A., LOBO, A., MAGNUSSON, H., MANN, A. H., MELLER, I., PRINCE,

- M. J., REISCHIES, F., TURRINA, C., DEVRIES, M. W., AND WILSON, K. C. **Depression in Europe. Geographical distribution among older people.** *The British journal of psychiatry : the journal of mental science* 174 (1999), 312–321.
- [24] WORLD HEALTH ORGANIZATION. **Depression and Other Common Mental Disorders.** World Health Organization, 2017.
- [25] LEMOGNE, C., PIOLINO, P., JOUVENT, R., ALLILAIRE, J.-F., AND FOSSATI, P. **Episodic autobiographical memory in depression: a review.** *L'Encephale* 32, 5 Pt 1 (2006), 781–788.
- [26] MARVEL, C. L., AND PARADISO, S. **Cognitive and neurological impairment in mood disorders.** *The Psychiatric clinics of North America* 27, 1 (2004), 19–36.
- [27] CHRISTENSEN, H., K., G., MACKINNON, A., AND JACOMB, P. **A quantitative review of cognitive deficits in depression and Alzheimer-type dementia.** *Journal of the International Neuropsychological Society* 3, 6 (1997), 631–651.
- [28] VEIEL, H. O. **A preliminary profile of neuropsychological deficits associated with major depression.** *Journal of clinical and experimental neuropsychology* 19, 4 (1997), 587–603.
- [29] ZAKZANIS, K. K., LEACH, L., AND KAPLAN, E. **On the nature and pattern of neurocognitive function in major depressive disorder.** *Neuropsychiatry, neuropsychology, and behavioral neurology* 11 (1998), 111–119.
- [30] RABINS, P. V., MERCHANT, A., AND NESTADT, G. **Criteria for diagnosing reversible dementia caused by depression: validation by 2-year follow-up.** *The British journal of psychiatry : the journal of mental science* 144 (1984), 488–492.
- [31] WORLD HEALTH ORGANIZATION. **Classification internationale des troubles mentaux et des troubles du comportement (cim-11/icd-11).** Elsevier/Masson, 2019.
- [32] FOLSTEIN, M. F., FOLSTEIN, S. E., AND MCHUGH, P. R. **"Mini-mental state". A practical method for grading the cognitive state of patients for the clinician.** *Journal of Psychiatric Research* 12, 3 (1975), 189–98.
- [33] DONNET, A., AND AZORIN, J. **L'année g erontologique.** Serdi, Paris, 1992, ch. Les pseudo-d emences d epressives., pp. 351–357.
- [34] ZAPOTOCZKY, H. G. **Problems of differential diagnosis between depressive pseudodementia and Alzheimer's disease.** Springer, Vienna, 1998.
- [35] ALEXOPOULOS, G. S., MEYERS, B. S., YOUNG, R. C., KALAYAM, B., KAKUMA, T., GABRIELLE, M., SIREY, J. A., AND HULL, J. **Executive dysfunction and long-term outcomes of geriatric depression.** *Archives of General Psychiatry* 57, 3 (2000), 285.
- [36] LYKETSOS, C. G., BREITNER, J. C., AND RABINS, P. V. **An evidence-based proposal for the classification of neuropsychiatric disturbance in Alzheimer's disease.** *International Journal of Geriatric Psychiatry* 16, 11 (2001), 1037–1042.

- [37] DANION, J. M., WILLARD-SCHROEDER, D., ZIMMERMANN, M. A., GRANGÉ, D., SCHLIENGER, J. L., AND SINGER, L. **Explicit memory and repetition priming in depression. Preliminary findings.** *Archives of General Psychiatry* 48, 8 (1991), 707–11.
- [38] PENN, E., AND TRACY, D. K. **The drugs don't work? Antidepressants and the current and future pharmacological management of depression.** *Therapeutic Advances in Psychopharmacology* 2, 5 (2012), 179–188.
- [39] AMIEVA, H., ANDRIEU, S., BERR, C., BUÉE, L., CHECLER, F., CLÉMENT, S., AND DARTIGUES, J.-F. **Maladie d'Alzheimer Enjeux scientifiques, Expertise collective.** INSERM, 2007.
- [40] SWAINSON, R., HODGES, J. R., GALTON, C. J., SEMPLE, J., MICHAEL, A., DUNN, B. D., IDDON, J. L., ROBBINS, T. W., AND SAHAKIAN, B. J. **Early detection and differential diagnosis of Alzheimer's disease and depression with neuropsychological tasks.** *Dementia and geriatric cognitive disorders* 12, 4 (2001), 265–280.
- [41] GALLARDA, T., AND LÔO, H. **Dépression et personnes âgées.** *L'Encéphale* 35, 3 (2009), 269–280.
- [42] LÔO, H., AND GALLARDA, T. **Troubles dépressifs et personnes âgées.** John Libbey Eurotext, 2000.
- [43] LAUREYS, S., AND BOLY, M. **What is it like to be vegetative or minimally conscious?** *Current Opinion in Neurology* 20, 6 (2007), 609–613.
- [44] JENNETT, B. **Thirty years of the vegetative state: clinical, ethical and legal problems.** *Progress in Brain Research* 150 (2005), 537–543.
- [45] POSNER, J. B., SAPER, C. B., SCHIFF, N., AND PLUM, F. **Diagnosis of Stupor and Coma.** Oxford University Press, 1966.
- [46] BRUNO, M. A., LEDOUX, D., VANHAUDENHUYSE, A., GOSSERIES, O., THIBAUT, A., AND LAUREYS, S. **Pronostic des patients récupérant du coma.** In *Coma et états de conscience altérée*, C. Schnakers and S. Laureys, Eds. Springer, Paris, 2011, p. 18.
- [47] AD-HOC COMMITTEE. **A Definition of Irreversible Coma.** Tech. Rep. 6, Harvard Medical School, 1968.
- [48] KING, J. S., AND MOULTON, B. W. **Rethinking informed consent: the case for shared medical decision-making.** *American journal of law and medicine* 32, 4 (2006), 429–501.
- [49] ZEMAN, A. **Consciousness: concepts, neurobiology, terminology of impairments, theoretical models and philosophical background.** In *Handbook of clinical neurology*, G. Young and E. Wijdicks, Eds., vol. 90. Elsevier B.V., 2008.
- [50] TEASDALE, G., AND JENNETT, B. **Assessment of coma and impaired consciousness. A practical scale.** *Lancet* 2, 7872 (1974), 81–84.
- [51] JAGGER, J., JANE, J., AND RIMEL, R. **The Glasgow coma scale: to sum or not to sum?** *The Lancet* 322, 8341 (1983), 97.

- [52] BHATTY, G. B., AND KAPOOR, N. **The Glasgow Coma Scale: a mathematical critique.** *Acta Neurochirurgica* 120, 3-4 (1993), 132–135.
- [53] GIACINO, J. T., KALMAR, K., AND WHYTE, J. **The JFK Coma Recovery Scale-Revised: measurement characteristics and diagnostic utility.** *Archives of physical medicine and rehabilitation* 85, 12 (2004), 2020–9.
- [54] SCHNAKERS, C. **Altération et récupération de la conscience chez les patients cérébro-lésés : Une approche comportementale, électrophysiologique et par neuro-imagerie fonctionnelle.** PhD thesis, Université de Liège, 2008.
- [55] SCHNAKERS, C., VANHAUDENHUYSE, A., GIACINO, J., VENTURA, M., BOLY, M., MAJERUS, S., MOONEN, G., AND LAUREYS, S. **Diagnostic accuracy of the vegetative and minimally conscious state: clinical consensus versus standardized neurobehavioral assessment.** *BMC neurology* 9 (2009), 35.
- [56] BIRKENHEAD, C. O. **Richard Caton (1842-1926) Pioneer Electrophysiologist.** *Proceedings of the Royal Society of Medicine* 52 (1959), 645–651.
- [57] BERGER, H. **Über das Elektrenkephalogramm des Menschen.** *Archiv für Psychiatrie und Nervenkrankheiten* 87, 1 (1929), 527–570.
- [58] DAWSON, G. **A summation technique for the detection of small evoked potentials.** *Electroencephalography and Clinical Neurophysiology* 6 (1954), 65–84.
- [59] QUIAN QUIROGA, R., SAKOWITZ, O., BASAR, E., AND SCHÜRMAN, M. **Wavelet transform in the analysis of the frequency composition of evoked potentials.** *Brain Research Protocols* 8, 1 (2001), 16–24.
- [60] MAJUMDAR, N., PRIBRAM, K., AND BARRETT, T. **Time frequency characterization of evoked brain activity in multiple electrode recordings.** *IEEE Transactions on Biomedical Engineering* 53, 12 (2006), 2516–2524.
- [61] DEMOTT, D. W. **Toposcopic studies of learning.** CC Thomas, Springfield, 1970.
- [62] LEHMANN, D. **Multichannel topography of human alpha EEG fields.** *Electroencephalography and Clinical Neurophysiology* 31, 5 (1971), 439–449.
- [63] LIVANOV, M. N. **Spatial organization of cerebral processes.** Wiley, New York, 1977.
- [64] MAURER, K. **Topographic brain mapping of EEG and evoked potentials.** Springer, Berlin, 1989.
- [65] FREEMAN, W. J., AND ROGERS, L. J. **Fine temporal resolution of analytic phase reveals episodic synchronization by state transitions in gamma EEGs.** *Journal of Neurophysiology* 87, 2 (2002), 937–945.
- [66] NIEDERMEYER, E., AND LOPES DA SILVA, F. **Electroencephalography: basic principles, clinical applications, and related fields**, fifth edition ed. Lippincott Williams & Wilkins, 2005.
- [67] KANDEL, E. R. **Principles of Neural Science**, fourth edition ed. McGraw-Hill, New York, 2000.

- [68] ZSCHOCKE, S. **Klinische Elektroenzephalographie**. Springer, Berlin, 1995.
- [69] NUNEZ, P. L., AND SRINIVASAN, R. **Electric Fields of the Brain**, second edition ed. Oxford University Press, New York, 2006.
- [70] D., L. **Handbook of Electroencephalography and Clinical Neurophysiology. Vol. 1: Methods of Analysis of Brain Electrical and Magnetic Signals**. Amsterdam: Elsevier, 1987, ch. Principles of spatial analysis, pp. 309–354.
- [71] BRANDEIS, D., AND LEHMANN, D. **Event-related potentials of the brain and cognitive processes: Approaches and applications**. *Neuropsychologia* 24, 1 (1986), 151–168.
- [72] LEHMANN, D., AND SKRANDIES, W. **Reference-free identification of components of checkerboard-evoked multichannel potential fields**. *Electroencephalography and Clinical Neurophysiology* 48, 6 (1980), 609–621.
- [73] SHAH, A. S. **Neural dynamics and the fundamental mechanisms of event-related brain potentials**. *Cerebral Cortex* 14, 5 (2004), 476–483.
- [74] MURAKAMI, S., AND OKADA, Y. **Contributions of principal neocortical neurons to magnetoencephalography and electroencephalography signals**. *The Journal of Physiology* 575, 3 (2006), 925–936.
- [75] SIULY, WANG, H., AND ZHANG, Y. **Detection of motor imagery EEG signals employing Naïve Bayes based learning process**. *Measurement* 86 (2016), 148–158.
- [76] NÓ, R. L. D. **Action potential of the motoneurons of the hypoglossus nucleus**. *Journal of Cellular and Comparative Physiology* 29, 3 (1947), 207–287.
- [77] ATWOOD, H. L., AND MACKAY, W. A. **Essentials of neurophysiology**. BC Decker, Toronto, 1989.
- [78] VIDAL, F. **Les interfaces cerveau-ordinateur 1**. iSTE editions, 2016, ch. Électrogenèse cérébrale, pp. 69–85.
- [79] ELUL, R. **The genesis of the EEG**. In *International Review of Neurobiology*. Elsevier, 1972, pp. 227–272.
- [80] CARLSON, N. R. **Foundations of physiological psychology**, fifth edition ed. Allyn and Bacon, Boston, 2002.
- [81] SANEI, S., AND CHAMBERS, J. **EEG signal processing**. Wiley, New York, 2007.
- [82] PURVES, D., AUGUSTINE, G. J., FITZPATRICK, D., HALL, W. C., LAMANTIA, A.-S., MOONEY, R. D., PLATT, M. L., AND WHITE, L. E. **Neurosciences**, sixième édition ed. De Boeck Supérieur, 2019.
- [83] CLYNES, M. **Dynamics of vertex evoked potentials: The R-M brain function**. In *Average evoked potentials: Methods, results, and evaluations*. US National Aeronautics and Space Administration, 1969, pp. 363–374.
- [84] PURHONEN, M., PÄÄKKÖNEN, A., YPPÄRILÄ, H., LEHTONEN, J., AND KARHU, J. **Dynamic behavior of the auditory N100 elicited by a baby's cry**. *International Journal of Psychophysiology* 41, 3 (2001), 271–278.

- [85] ALHO, K. **Cerebral generators of mismatch negativity (MMN) and its magnetic counterpart (MMNm) elicited by sound changes.** *Ear and Hearing* 16, 1 (1995), 38–51.
- [86] NÄÄTÄNEN, R., PAAVILAINEN, P., RINNE, T., AND ALHO, K. **The mismatch negativity (MMN) in basic research of central auditory processing: A review.** *Clinical Neurophysiology* 118, 12 (2007), 2544–2590.
- [87] COURCHESNE, E., HILLYARD, S. A., AND GALAMBOS, R. **Stimulus novelty, task relevance and the visual evoked potential in man.** *Electroencephalography and Clinical Neurophysiology* 39, 2 (1975), 131–143.
- [88] COLES, M., AND RUGG, M. **Event-related brain potentials: An introduction.** In *Electrophysiology of Mind: Event-Related Brain Potentials and Cognition*, M. Rugg and M. Coles, Eds. Oxford University Press, London, 1995, pp. 1–26.
- [89] POLICH, J. **Meta-analysis of P300 normative aging studies.** *Psychophysiology* 33, 4 (1996), 334–353.
- [90] KUTAS, M., AND HILLYARD, S. A. **Event-related brain potentials to semantically inappropriate and surprisingly large words.** *Biological Psychology* 11, 2 (1980), 99–116.
- [91] CHAO, L., NIELSEN-BOHLMAN, L., AND KNIGHT, R. **Auditory event-related potentials dissociate early and late memory processes.** *Electroencephalography and Clinical Neurophysiology/Evoked Potentials Section* 96, 2 (1995), 157–168.
- [92] VANPETTEN, C., AND RHEINFELDER, H. **Conceptual relationships between spoken words and environmental sounds: Event-related brain potential measures.** *Neuropsychologia* 33, 4 (1995), 485–508.
- [93] CONNOLLY, J. **The effects of processing requirements on neurophysiological responses to spoken sentences.** *Brain and Language* 39, 2 (1990), 302–318.
- [94] PERRIN, F., MAQUET, P., PEIGNEUX, P., RUBY, P., DEGUELDRE, C., BALTEAU, E., FIORE, G. D., MOONEN, G., LUXEN, A., AND LAUREYS, S. **Neural mechanisms involved in the detection of our first name: a combined ERPs and PET study.** *Neuropsychologia* 43, 1 (2005), 12–19.
- [95] MUTHUSWAMY, J., SRIDHARAN, A., AND OKANDAN, M. **MEMS Neural Probes.** In *Encyclopedia of Nanotechnology*, B. Bhushan, Ed. Springer, Dordrecht, 2016.
- [96] KUTLU, Y., KUNTALP, M., AND KUNTALP, D. **Optimizing the performance of an MLP classifier for the automatic detection of epileptic spikes.** *Expert Systems with Applications* 36, 4 (2009), 7567–7575.
- [97] SUBASI, A., AND ERÇELEBI, E. **Classification of EEG signals using neural network and logistic regression.** *Computer Methods and Programs in Biomedicine* 78, 2 (2005), 87–99.
- [98] SABOURET, N. **Comprendre l'intelligence artificielle.** Ellipses, Paris, 2019.
- [99] JAIN, A., DUIN, P., AND MAO, J. **Statistical pattern recognition: a review.** *IEEE Transactions on Pattern Analysis and Machine Intelligence* 22, 1 (2000), 4–37.

- [100] DUDA, R. O., HART, P. E., AND STORK, D. G. **Pattern Classification**. Wiley John + Sons, 2000.
- [101] BRUNELLI, R. **Template Matching Techniques in Computer Vision: Theory and Practice**. WILEY, 2009.
- [102] COVER, T., AND HART, P. **Nearest neighbor pattern classification**. *IEEE Transactions on Information Theory* 13, 1 (1967), 21–27.
- [103] HAN, J., KAMBER, M., AND PEI, J. **Data Mining**. Elsevier, 2012.
- [104] RIPLEY, B. D. **Pattern Recognition and Neural Networks**. Cambridge University Press, 2012.
- [105] FRIEDMAN, N., GEIGER, D., AND GOLDSZMIDT, M. **Bayesian network classifiers**. *Machine Learning* 29, 2/3 (1997), 131–163.
- [106] VAPNIK, V. **Pattern recognition using generalized portrait method**. *Automation and Remote Control* 24 (1963), 774–780.
- [107] HERBRICH, R. **Learning Kernel Classifiers: Theory and Algorithms**. MIT Press, Cambridge, MA, USA, 2001.
- [108] CHAWLA, N. V., BOWYER, K. W., HALL, L. O., AND KEGELMEYER, W. P. **SMOTE: Synthetic minority over-sampling technique**. *Journal of Artificial Intelligence Research* 16 (2002), 321–357.
- [109] DAS, K., GIESBRECHT, B., AND ECKSTEIN, M. P. **Predicting variations of perceptual performance across individuals from neural activity using pattern classifiers**. *NeuroImage* 51, 4 (2010), 1425–1437.
- [110] ACHARYA, U. R., SUDARSHAN, V. K., ADELI, H., SANTHOSH, J., KOH, J. E., PUTHANKATTI, S. D., AND ADELI, A. **A Novel Depression Diagnosis Index Using Nonlinear Features in EEG Signals**. *European Neurology* 74, 1-2 (2015), 79–83.
- [111] KAKKOS, I., VENTOURAS, E. M., ASVESTAS, P. A., KARANASIOU, I. S., AND MATSOPOULOS, G. K. **A condition-independent framework for the classification of error-related brain activity**. *Medical and Biological Engineering and Computing* 58, 3 (2020), 573–587.
- [112] WANG, R., WANG, J., YU, H., WEI, X., YANG, C., AND DENG, B. **Power spectral density and coherence analysis of Alzheimer’s EEG**. *Cognitive Neurodynamics* 9, 3 (2015), 291–304.
- [113] PADILLA-BURITICA, J. I., MARTINEZ-VARGAS, J. D., AND CASTELLANOS-DOMINGUEZ, G. **Emotion Discrimination Using Spatially Compact Regions of Interest Extracted from Imaging EEG Activity**. *Frontiers in Computational Neuroscience* 10 (2016).
- [114] AL-KAYSI, A. M., AL-ANI, A., LOO, C. K., POWELL, T. Y., MARTIN, D. M., BREAKSPEAR, M., AND BOONSTRA, T. W. **Predicting tDCS treatment outcomes of patients with major depressive disorder using automated EEG classification**. *Journal of Affective Disorders* 208 (2017), 597–603.

- [115] ALAZRAI, R., HOMOUD, R., ALWANNI, H., AND DAOUD, M. **EEG-Based Emotion Recognition Using Quadratic Time-Frequency Distribution**. *Sensors* 18, 8 (2018), 2739.
- [116] TIAN, Y., ZHANG, H., JIANG, Y., LI, P., AND LI, Y. **A Fusion Feature for Enhancing the Performance of Classification in Working Memory Load With Single-Trial Detection**. *IEEE Transactions on Neural Systems and Rehabilitation Engineering* 27, 10 (2019), 1985–1993.
- [117] MOHSENI, M., SHALCHYAN, V., JOCHUMSEN, M., AND NIAZI, I. K. **Upper limb complex movements decoding from pre-movement EEG signals using wavelet common spatial patterns**. *Computer Methods and Programs in Biomedicine* 183 (2020), 105076.
- [118] RIVET, B., SOULOUMIAC, A., ATTINA, V., AND GIBERT, G. **xDAWN algorithm to enhance evoked potentials: Application to brain–computer interface**. *IEEE Transactions on Biomedical Engineering* 56, 8 (2009), 2035–2043.
- [119] BASHASHATI, A., FATOURECHI, M., WARD, R. K., AND BIRCH, G. E. **A survey of signal processing algorithms in brain–computer interfaces based on electrical brain signals**. *Journal of Neural Engineering* 4, 2 (2007), R32–R57.
- [120] POLAT, K., AND GÜNEŞ, S. **Classification of epileptiform EEG using a hybrid system based on decision tree classifier and fast fourier transform**. *Applied Mathematics and Computation* 187, 2 (2007), 1017–1026.
- [121] QIN, L., AND HE, B. **A wavelet-based time–frequency analysis approach for classification of motor imagery for brain–computer interface applications**. *Journal of Neural Engineering* 2, 4 (2005), 65–72.
- [122] TING, W., GUO-ZHENG, Y., BANG-HUA, Y., AND HONG, S. **EEG feature extraction based on wavelet packet decomposition for brain computer interface**. *Measurement* 41, 6 (2008), 618–625.
- [123] LOTTE, F., AND GUAN, C. **Regularizing common spatial patterns to improve BCI designs: Unified theory and new algorithms**. *IEEE Transactions on Biomedical Engineering* 58, 2 (2011), 355–362.
- [124] LU, H., ENG, H.-L., GUAN, C., PLATANIOTIS, K. N., AND VENETSANOPOULOS, A. N. **Regularized common spatial pattern with aggregation for EEG classification in small-sample setting**. *IEEE Transactions on Biomedical Engineering* 57, 12 (2010), 2936–2946.
- [125] ZHANG, R., XU, P., GUO, L., ZHANG, Y., LI, P., AND YAO, D. **Z-score linear discriminant analysis for EEG based brain-computer interfaces**. *PLoS ONE* 8, 9 (2013), e74433.
- [126] GUO, J., GAO, S., AND HONG, B. **An Auditory Brain-Computer Interface Using Active Mental Response**. *IEEE Transactions on Neural Systems and Rehabilitation Engineering* 18, 3 (2010), 230–235.
- [127] YANG, T., AND KIM, S.-P. **Group-Level Neural Responses to Service-to-Service Brand Extension**. *Frontiers in Neuroscience* 13 (2019), 676.

- [128] SIULY, LI, Y., AND WEN, P. P. **Clustering technique-based least square support vector machine for EEG signal classification.** *Computer Methods and Programs in Biomedicine* 104, 3 (2011), 358–372.
- [129] GUO, L., RIVERO, D., SEOANE, J. A., AND PAZOS, A. **Classification of EEG signals using relative wavelet energy and artificial neural networks.** In *Proceedings of the first ACM/SIGEVO Summit on Genetic and Evolutionary Computation - GEC '09* (2009), ACM Press.
- [130] SIULY, S., AND LI, Y. **Improving the separability of motor imagery EEG signals using a cross correlation-based least square support vector machine for brain–computer interface.** *IEEE Transactions on Neural Systems and Rehabilitation Engineering* 20, 4 (2012), 526–538.
- [131] SIULY, S., AND LI, Y. **Discriminating the brain activities for brain–computer interface applications through the optimal allocation-based approach.** *Neural Computing and Applications* 26, 4 (2014), 799–811.
- [132] FAWCETT, T. **An introduction to ROC analysis.** *Pattern Recognition Letters* 27, 8 (2006), 861–874.
- [133] ARVANEH, M., GUAN, C., ANG, K. K., AND QUEK, C. **Optimizing the Channel Selection and Classification Accuracy in EEG-Based BCI.** *IEEE Transactions on Biomedical Engineering* 58, 6 (2011), 1865–1873.
- [134] KIRAR, J. S., AND AGRAWAL, R. K. **Relevant Feature Selection from a Combination of Spectral-Temporal and Spatial Features for Classification of Motor Imagery EEG.** *Journal of Medical Systems* 42, 5 (2018), 78.
- [135] HOU, H.-R., ZHANG, X.-N., AND MENG, Q.-H. **Odor-induced emotion recognition based on average frequency band division of EEG signals.** *Journal of Neuroscience Methods* 334 (2020), 108599.
- [136] ROY, Y., BANVILLE, H., ALBUQUERQUE, I., GRAMFORT, A., FALK, T. H., AND FAUBERT, J. **Deep learning-based electroencephalography analysis: a systematic review.** *Journal of Neural Engineering* 16, 5 (2019), 051001.
- [137] AGUIÑAGA, A. R., AND RAMIREZ, M. A. L. **Emotional states recognition, implementing a low computational complexity strategy.** *Health Informatics Journal* 24, 2 (2018), 146–170.
- [138] AL-SHARGIE, F., TANG, T. B., BADRUDDIN, N., AND KIGUCHI, M. **Towards multilevel mental stress assessment using SVM with ECOC: an EEG approach.** *Medical and Biological Engineering and Computing* 56, 1 (2018), 125–136.
- [139] AMIN, H. U., MUMTAZ, W., SUBHANI, A. R., SAAD, M. N. M., AND MALIK, A. S. **Classification of EEG Signals Based on Pattern Recognition Approach.** *Frontiers in Computational Neuroscience* 11 (2017), 103.
- [140] AREVALILLO-HERRÁEZ, M., COBOS, M., ROGER, S., AND GARCÍA-PINEDA, M. **Combining Inter-Subject Modeling with a Subject-Based Data Transformation to Improve Affect Recognition from EEG Signals.** *Sensors* 19, 13 (2019), 2999.

- [141] ARSALAN, A., MAJID, M., BUTT, A. R., AND ANWAR, S. M. **Classification of Perceived Mental Stress Using A Commercially Available EEG Headband.** *IEEE Journal of Biomedical and Health Informatics* 23, 6 (2019), 2257–2264.
- [142] BĂLAN, O., MOISE, G., MOLDOVEANU, A., LEORDEANU, M., AND MOLDOVEANU, F. **Fear Level Classification Based on Emotional Dimensions and Machine Learning Techniques.** *Sensors* 19, 7 (2019), 1738.
- [143] IACOVIELLO, D., PETRACCA, A., SPEZIALETTI, M., AND PLACIDI, G. **A real-time classification algorithm for EEG-based BCI driven by self-induced emotions.** *Computer Methods and Programs in Biomedicine* 122, 3 (2015), 293–303.
- [144] IACOVIELLO, D., PETRACCA, A., SPEZIALETTI, M., AND PLACIDI, G. **A Classification Algorithm for Electroencephalography Signals by Self-Induced Emotional Stimuli.** *IEEE Transactions on Cybernetics* 46, 12 (2016), 3171–3180.
- [145] JAHROMY, F. Z., BAJOULVAND, A., AND DALIRI, M. R. **Statistical algorithms for emotion classification via functional connectivity.** *Journal of Integrative Neuroscience* 18, 3 (2019), 293.
- [146] LIN, Y.-P., WANG, C.-H., JUNG, T.-P., WU, T.-L., JENG, S.-K., DUANN, J.-R., AND CHEN, J.-H. **EEG-Based Emotion Recognition in Music Listening.** *IEEE Transactions on Biomedical Engineering* 57, 7 (2010), 1798–1806.
- [147] LIU, S., ZHANG, D., XU, M., QI, H., HE, F., ZHAO, X., ZHOU, P., ZHANG, L., AND MING, D. **Randomly dividing homologous samples leads to overinflated accuracies for emotion recognition.** *International Journal of Psychophysiology* 96, 1 (2015), 29–37.
- [148] LIU, S., CHEN, L., GUO, D., LIU, X., SHENG, Y., KE, Y., XU, M., AN, X., YANG, J., AND MING, D. **Incorporation of Multiple-Days Information to Improve the Generalization of EEG-Based Emotion Recognition Over Time.** *Frontiers in Human Neuroscience* 12 (2018), 267.
- [149] LIU, X., WANG, Q., LIU, D., WANG, Y., ZHANG, Y., BAI, O., AND SUN, J. **Human emotion classification based on multiple physiological signals by wearable system.** *Technology and Health Care* 26 (2018), 459–469.
- [150] MUUKKONEN, I., ÖLANDER, K., NUMMINEN, J., AND SALMELA, V. **Spatio-temporal dynamics of face perception.** *NeuroImage* 209 (2020), 116531.
- [151] PETRANTONAKIS, P., AND HADJILEONTIADIS, L. **Emotion Recognition From EEG Using Higher Order Crossings.** *IEEE Transactions on Information Technology in Biomedicine* 14, 2 (2010), 186–197.
- [152] RAHEEL, A., MAJID, M., AND ANWAR, S. M. **A study on the effects of traditional and olfaction enhanced multimedia on pleasantness classification based on brain activity analysis.** *Computers in Biology and Medicine* 114 (2019), 103469.
- [153] RAMZAN, M., AND DAWN, S. **Learning-based classification of valence emotion from electroencephalography.** *International Journal of Neuroscience* 129, 11 (2019), 1085–1093.

- [154] SORINAS, J., GRIMA, M. D., FERRANDEZ, J. M., AND FERNANDEZ, E. **Identifying Suitable Brain Regions and Trial Size Segmentation for Positive/Negative Emotion Recognition.** *International Journal of Neural Systems* 29, 02 (2019), 1850044.
- [155] TARAN, S., AND BAJAJ, V. **Emotion recognition from single-channel EEG signals using a two-stage correlation and instantaneous frequency-based filtering method.** *Computer Methods and Programs in Biomedicine* 173 (2019), 157–165.
- [156] THAMMASAN, N., MORIYAMA, K., FUKUI, K.-I., AND NUMAO, M. **Familiarity effects in EEG-based emotion recognition.** *Brain Informatics* 4, 1 (2017), 39–50.
- [157] WEI, W., JIA, Q., FENG, Y., AND CHEN, G. **Emotion Recognition Based on Weighted Fusion Strategy of Multichannel Physiological Signals.** *Computational Intelligence and Neuroscience* 2018 (2018), 1–9.
- [158] YIN, Z., WANG, Y., LIU, L., ZHANG, W., AND ZHANG, J. **Cross-Subject EEG Feature Selection for Emotion Recognition Using Transfer Recursive Feature Elimination.** *Frontiers in Neurobotics* 11 (2017).
- [159] AYDIN, S., GÜDÜCÜ, C., KUTLUK, F., ÖNİZ, A., AND ÖZGÖREN, M. **The impact of musical experience on neural sound encoding performance.** *Neuroscience Letters* 694 (2019), 124–128.
- [160] BAMATRAF, S., HUSSAIN, M., ABOALSAMH, H., QAZI, E.-U.-H., MALIK, A. S., AMIN, H. U., MATHKOUR, H., MUHAMMAD, G., AND IMRAN, H. M. **A System for True and False Memory Prediction Based on 2D and 3D Educational Contents and EEG Brain Signals.** *Computational Intelligence and Neuroscience* 2016 (2016), 1–11.
- [161] BATBAT, T., GÜVEN, A., AND DOLU, N. **Evaluation of divided attention using different stimulation models in event-related potentials.** *Medical and Biological Engineering and Computing* 57, 9 (2019), 2069–2079.
- [162] BINIAS, B., MYZOR, D., AND CYRAN, K. A. **A Machine Learning Approach to the Detection of Pilot’s Reaction to Unexpected Events Based on EEG Signals.** *Computational Intelligence and Neuroscience* 2018 (2018), 1–9.
- [163] BOCINCOVA, A., AND JOHNSON, J. S. **The time course of encoding and maintenance of task-relevant versus irrelevant object features in working memory.** *Cortex* 111 (2019), 196–209.
- [164] BOJORGES-VALDEZ, E., ECHEVERRÍA, J. C., AND YANEZ-SUAREZ, O. **Evaluation of the continuous detection of mental calculation episodes as a BCI control input.** *Computers in Biology and Medicine* 64 (2015), 155–162.
- [165] BOSCH, P., HERRERA, M., LÓPEZ, J., AND MALDONADO, S. **Mining EEG with SVM for Understanding Cognitive Underpinnings of Math Problem Solving Strategies.** *Behavioural Neurology* 2018 (2018), 1–15.
- [166] CHEN, J., WANG, H., AND HUA, C. **Assessment of driver drowsiness using electroencephalogram signals based on multiple functional brain networks.** *International Journal of Psychophysiology* 133 (2018), 120–130.

- [167] CHEN, J., WANG, H., WANG, Q., AND HUA, C. **Exploring the fatigue affecting electroencephalography based functional brain networks during real driving in young males.** *Neuropsychologia* 129 (2019), 200–211.
- [168] CHOI, J. W., AND KIM, K. H. **Covert intention to answer “yes” or “no” can be decoded from single-trial electroencephalograms (EEGs).** *Computational Intelligence and Neuroscience* 2019 (2019), 1–11.
- [169] DONG, S.-Y., KIM, B.-K., AND LEE, S.-Y. **EEG-Based Classification of Implicit Intention During Self-Relevant Sentence Reading.** *IEEE Transactions on Cybernetics* 46, 11 (2016), 2535–2542.
- [170] GAO, J.-F., YANG, Y., HUANG, W.-T., LIN, P., GE, S., ZHENG, H.-M., GU, L.-Y., ZHOU, H., LI, C.-H., AND RAO, N.-N. **Exploring time- and frequency- dependent functional connectivity and brain networks during deception with single-trial event-related potentials.** *Scientific Reports* 6, 1 (2016), 37065.
- [171] GAO, Z.-K., GUO, W., CAI, Q., MA, C., ZHANG, Y.-B., AND KURTHS, J. **Characterization of SSMVEP-based EEG signals using multiplex limited penetrable horizontal visibility graph.** *Chaos: An Interdisciplinary Journal of Nonlinear Science* 29, 7 (2019), 073119.
- [172] GURSEL OZMEN, N., GUMUSEL, L., AND YANG, Y. **A Biologically Inspired Approach to Frequency Domain Feature Extraction for EEG Classification.** *Computational and Mathematical Methods in Medicine* 2018 (2018), 1–10.
- [173] HU, J., AND MIN, J. **Automated detection of driver fatigue based on EEG signals using gradient boosting decision tree model.** *Cognitive Neurodynamics* 12, 4 (2018), 431–440.
- [174] KE, Y., QI, H., ZHANG, L., CHEN, S., JIAO, X., ZHOU, P., ZHAO, X., WAN, B., AND MING, D. **Towards an effective cross-task mental workload recognition model using electroencephalography based on feature selection and support vector machine regression.** *International Journal of Psychophysiology* 98, 2 (2015), 157–166.
- [175] KRUMPE, T., SCHARINGER, C., ROSENSTIEL, W., GERJETS, P., AND SPÜLER, M. **Unity and diversity in working memory load: Evidence for the separability of the executive functions updating and inhibition using machine learning.** *Biological Psychology* 139 (2018), 163–172.
- [176] KUO, C.-H., CHEN, H.-H., CHOU, H.-C., CHEN, P.-N., AND KUO, Y.-C. **Wireless Stimulus-on-Device Design for Novel P300 Hybrid Brain-Computer Interface Applications.** *Computational Intelligence and Neuroscience* 2018 (2018), 1–13.
- [177] LIANG, Z., SHAO, S., LV, Z., LI, D., SLEIGH, J. W., LI, X., ZHANG, C., AND HE, J. **Constructing a Consciousness Meter Based on the Combination of Non-Linear Measurements and Genetic Algorithm-Based Support Vector Machine.** *IEEE Transactions on Neural Systems and Rehabilitation Engineering* 28, 2 (2020), 399–408.

- [178] MOHAN, D. M., KUMAR, P., MAHMOOD, F., WONG, K. F., AGRAWAL, A., ELGENDI, M., SHUKLA, R., ANG, N., CHING, A., DAUWELS, J., AND CHAN, A. H. D. **Effect of Subliminal Lexical Priming on the Subjective Perception of Images: A Machine Learning Approach.** *PLOS ONE* 11, 2 (2016), e0148332.
- [179] PLEWAN, T., WASCHER, E., FALKENSTEIN, M., AND HOFFMANN, S. **Classifying Response Correctness across Different Task Sets: A Machine Learning Approach.** *PLOS ONE* 11, 3 (2016), e0152864.
- [180] RAHMAN, M. M., CHOWDHURY, M. A., AND FATTAH, S. A. **An efficient scheme for mental task classification utilizing reflection coefficients obtained from autocorrelation function of EEG signal.** *Brain Informatics* 5, 1 (2018), 1–12.
- [181] RAHMAN, M. M., AND FATTAH, S. A. **Mental Task Classification Scheme Utilizing Correlation Coefficient Extracted from Interchannel Intrinsic Mode Function.** *BioMed Research International 2017* (2017), 1–11.
- [182] SERESHKEH, A. R., TROTT, R., BRICOUT, A., AND CHAU, T. **Online EEG Classification of Covert Speech for Brain-Computer Interfacing.** *International Journal of Neural Systems* 27, 08 (2017), 1750033.
- [183] SHALBAF, A., SHALBAF, R., SAFFAR, M., AND SLEIGH, J. **Monitoring the level of hypnosis using a hierarchical SVM system.** *Journal of Clinical Monitoring and Computing* 34, 2 (2020), 331–338.
- [184] STERN, Y., RECHES, A., AND GEVA, A. B. **Brain Network Activation Analysis Utilizing Spatiotemporal Features for Event Related Potentials Classification.** *Frontiers in Computational Neuroscience* 10 (2016).
- [185] SWEETI, JOSHI, D., PANIGRAHI, B. K., ANAND, S., AND SANTHOSH, J. **Classification of Targets and Distractors Present in Visual Hemifields Using Time-Frequency Domain EEG Features.** *Journal of Healthcare Engineering 2018* (2018), 1–10.
- [186] TAFRESHI, T. F., DALIRI, M. R., AND GHODOUSI, M. **Functional and effective connectivity based features of EEG signals for object recognition.** *Cognitive Neurodynamics* 13, 6 (2019), 555–566.
- [187] TIAN, Y., ZHANG, H., PANG, Y., AND LIN, J. **Classification for Single-Trial N170 During Responding to Facial Picture With Emotion.** *Frontiers in Computational Neuroscience* 12 (2018), 68.
- [188] TIAN, Y., XU, W., AND YANG, L. **Cortical Classification with Rhythm Entropy for Error Processing in Cocktail Party Environment Based on Scalp EEG Recording.** *Scientific Reports* 8, 1 (2018), 6070.
- [189] VANDENBOSCH, M. M. L. J. Z., VAN 'T ENT, D., BOOMSMA, D. I., ANOKHIN, A. P., AND SMIT, D. J. A. **EEG-based age-prediction models as stable and heritable indicators of brain maturational level in children and adolescents.** *Human Brain Mapping* 40, 6 (2019), 1919–1926.
- [190] WANG, Y., WANG, P., AND YU, Y. **Decoding English Alphabet Letters Using EEG Phase Information.** *Frontiers in Neuroscience* 12 (2018), 62.

- [191] WITTEVRONGEL, B., AND VAN HULLE, M. M. **Faster P300 Classifier Training Using Spatiotemporal Beamforming.** *International Journal of Neural Systems* 26, 03 (2016), 1650014.
- [192] YAN, F., LIU, M., DING, C., WANG, Y., AND YAN, L. **Driving Style Recognition Based on Electroencephalography Data From a Simulated Driving Experiment.** *Frontiers in Psychology* 10 (2019), 1254.
- [193] ZHANG, X., LI, J., LIU, Y., ZHANG, Z., WANG, Z., LUO, D., ZHOU, X., ZHU, M., SALMAN, W., HU, G., AND WANG, C. **Design of a Fatigue Detection System for High-Speed Trains Based on Driver Vigilance Using a Wireless Wearable EEG.** *Sensors* 17, 3 (2017), 486.
- [194] ANDRADE, J. A., CECÍLIO, J., SIMÕES, M., SALES, F., AND CASTELO-BRANCO, M. **Separability of motor imagery of the self from interpretation of motor intentions of others at the single trial level: an EEG study.** *Journal of NeuroEngineering and Rehabilitation* 14, 1 (2017), 63.
- [195] ARICAN, M., AND POLAT, K. **Pairwise and variance based signal compression algorithm (PVBSC) in the P300 based speller systems using EEG signals.** *Computer Methods and Programs in Biomedicine* 176 (2019), 149–157.
- [196] AYDEMIR, O., AND ERGÜN, E. **A robust and subject-specific sequential forward search method for effective channel selection in brain computer interfaces.** *Journal of Neuroscience Methods* 313 (2019), 60–67.
- [197] BASHASHATI, H., WARD, R. K., BIRCH, G. E., AND BASHASHATI, A. **Comparing Different Classifiers in Sensory Motor Brain Computer Interfaces.** *PLOS ONE* 10, 6 (2015), e0129435.
- [198] BIAN, Y., QI, H., ZHAO, L., MING, D., GUO, T., AND FU, X. **Improvements in event-related desynchronization and classification performance of motor imagery using instructive dynamic guidance and complex tasks.** *Computers in Biology and Medicine* 96 (2018), 266–273.
- [199] CHEN, M., GUAN, J., AND LIU, H. **Enabling Fast Brain-Computer Interaction by Single-Trial Extraction of Visual Evoked Potentials.** *Journal of Medical Systems* 35, 5 (2011), 1323–1331.
- [200] COSTA, Á., IÁÑEZ, E., ÚBEDA, A., HORTAL, E., DEL-AMA, A. J., GIL-AGUDO, Á., AND AZORÍN, J. M. **Decoding the Attentional Demands of Gait through EEG Gamma Band Features.** *PLOS ONE* 11, 4 (2016), e0154136.
- [201] DONG, E., LI, C., LI, L., DU, S., BELKACEM, A. N., AND CHEN, C. **Classification of multi-class motor imagery with a novel hierarchical SVM algorithm for brain-computer interfaces.** *Medical and Biological Engineering and Computing* 55, 10 (2017), 1809–1818.
- [202] HSU, W.-Y. **Continuous EEG signal analysis for asynchronous BCI application.** *International Journal of Neural Systems* 21, 04 (2011), 335–350.
- [203] ISLAM, M. R., TANAKA, T., AND MOLLA, M. K. I. **Multiband tangent space mapping and feature selection for classification of EEG during motor imagery.** *Journal of Neural Engineering* 15, 4 (2018), 046021.

- [204] JI, MA, DONG, AND ZHANG. **EEG Signals Feature Extraction Based on DWT and EMD Combined with Approximate Entropy.** *Brain Sciences* 9, 8 (2019), 201.
- [205] JIN, J., MIAO, Y., DALY, I., ZUO, C., HU, D., AND CICHOCKI, A. **Correlation-based channel selection and regularized feature optimization for MI-based BCI.** *Neural Networks* 118 (2019), 262–270.
- [206] JOADDER, M. A. M., MYSEWSKI, J. J., RAHMAN, M. H., AND WANG, I. **A performance based feature selection technique for subject independent MI based BCI.** *Health Information Science and Systems* 7, 1 (2019), 15.
- [207] JOCHUMSEN, M., KHAN NIAZI, I., TAYLOR, D., FARINA, D., AND DREMSTRUP, K. **Detecting and classifying movement-related cortical potentials associated with hand movements in healthy subjects and stroke patients from single-electrode, single-trial EEG.** *Journal of Neural Engineering* 12, 5 (2015), 056013.
- [208] JOCHUMSEN, M., NIAZI, I. K., MRACHACZ-KERSTING, N., JIANG, N., FARINA, D., AND DREMSTRUP, K. **Comparison of spatial filters and features for the detection and classification of movement-related cortical potentials in healthy individuals and stroke patients.** *Journal of Neural Engineering* 12, 5 (2015), 056003.
- [209] JOCHUMSEN, M., ROVSING, C., ROVSING, H., NIAZI, I. K., DREMSTRUP, K., AND KAMAVUAKO, E. N. **Classification of Hand Grasp Kinetics and Types Using Movement-Related Cortical Potentials and EEG Rhythms.** *Computational Intelligence and Neuroscience 2017* (2017), 1–8.
- [210] KETENCI, S., AND KAYIKCIOGLU, T. **Investigation of Theta Rhythm Effect in Detection of Finger Movement.** *Journal of Experimental Neuroscience* 13 (2019), 117906951982873.
- [211] KIM, H., YOSHIMURA, N., AND KOIKE, Y. **Characteristics of Kinematic Parameters in Decoding Intended Reaching Movements Using Electroencephalography (EEG).** *Frontiers in Neuroscience* 13 (2019), 1148.
- [212] KIM, H., YOSHIMURA, N., AND KOIKE, Y. **Classification of Movement Intention Using Independent Components of Premovement EEG.** *Frontiers in Human Neuroscience* 13 (2019), 63.
- [213] LI, Q., SHI, K., GAO, N., LI, J., AND BAI, O. **Training set extension for SVM ensemble in P300-speller with familiar face paradigm.** *Technology and Health Care* 26, 3 (2018), 469–482.
- [214] LI, M., WANG, R., YANG, J., AND DUAN, L. **An Improved Refined Composite Multivariate Multiscale Fuzzy Entropy Method for MI-EEG Feature Extraction.** *Computational Intelligence and Neuroscience 2019* (2019), 1–12.
- [215] LI, F., FAN, Y., ZHANG, X., WANG, C., HU, F., JIA, W., AND HUI, H. **Multi-Feature Fusion Method Based on EEG Signal and its Application in Stroke Classification.** *Journal of Medical Systems* 44, 2 (2020), 39.

- [216] LIU, W., LIU, X., DAI, R., AND TANG, X. **Exploring differences between left and right hand motor imagery via spatio-temporal EEG microstate.** *Computer Assisted Surgery* 22, sup1 (2017), 258–266.
- [217] LUDWIG, S. A., AND KONG, J. **Investigation of different classifiers and channel configurations of a mobile P300-based brain-computer interface.** *Medical and Biological Engineering and Computing* 55, 12 (2017), 2143–2154.
- [218] LUO, J., WANG, J., XU, R., AND XU, K. **Class discrepancy-guided sub-band filter-based common spatial pattern for motor imagery classification.** *Journal of Neuroscience Methods* 323 (2019), 98–107.
- [219] MA, Y., DING, X., SHE, Q., LUO, Z., POTTER, T., AND ZHANG, Y. **Classification of Motor Imagery EEG Signals with Support Vector Machines and Particle Swarm Optimization.** *Computational and Mathematical Methods in Medicine 2016* (2016), 1–8.
- [220] MALAN, N. S., AND SHARMA, S. **Feature selection using regularized neighbourhood component analysis to enhance the classification performance of motor imagery signals.** *Computers in Biology and Medicine* 107 (2019), 118–126.
- [221] MONDINI, V., MANGIA, A. L., AND CAPPELLO, A. **EEG-Based BCI System Using Adaptive Features Extraction and Classification Procedures.** *Computational Intelligence and Neuroscience 2016* (2016), 1–14.
- [222] PAN, J., XIE, Q., HE, Y., WANG, F., DI, H., LAUREYS, S., YU, R., AND LI, Y. **Detecting awareness in patients with disorders of consciousness using a hybrid brain-computer interface.** *Journal of Neural Engineering* 11, 5 (2014), 056007.
- [223] RATHEE, D., RAZA, H., PRASAD, G., AND CECOTTI, H. **Current Source Density Estimation Enhances the Performance of Motor-Imagery-Related Brain-Computer Interface.** *IEEE Transactions on Neural Systems and Rehabilitation Engineering* 25, 12 (2017), 2461–2471.
- [224] ROY, R., SIKDAR, D., MAHADEVAPPA, M., AND KUMAR, C. S. **A fingertip force prediction model for grasp patterns characterised from the chaotic behaviour of EEG.** *Medical and Biological Engineering and Computing* 56, 11 (2018), 2095–2107.
- [225] SALVARIS, M., AND SEPULVEDA, F. **Visual modifications on the P300 speller BCI paradigm.** *Journal of Neural Engineering* 6, 4 (2009), 046011.
- [226] SHAN, H., XU, H., ZHU, S., AND HE, B. **A novel channel selection method for optimal classification in different motor imagery BCI paradigms.** *BioMedical Engineering OnLine* 14, 1 (2015), 93.
- [227] SOUSA, T., AMARAL, C., ANDRADE, J., PIRES, G., NUNES, U. J., AND CASTELOBRANCO, M. **Pure visual imagery as a potential approach to achieve three classes of control for implementation of BCI in non-motor disorders.** *Journal of Neural Engineering* 14, 4 (2017), 046026.
- [228] SPÜLER, M., LÓPEZ-LARRAZ, E., AND RAMOS-MURGUIALDAY, A. **On the design of EEG-based movement decoders for completely paralyzed stroke patients.** *Journal of NeuroEngineering and Rehabilitation* 15, 1 (2018), 110.

- [229] TAVAKOLAN, M., FREHLICK, Z., YONG, X., AND MENON, C. **Classifying three imaginary states of the same upper extremity using time-domain features.** *PLOS ONE* 12, 3 (2017), e0174161.
- [230] YI, W., QIU, S., WANG, K., QI, H., HE, F., ZHOU, P., ZHANG, L., AND MING, D. **EEG oscillatory patterns and classification of sequential compound limb motor imagery.** *Journal of NeuroEngineering and Rehabilitation* 13, 1 (2016), 11.
- [231] ZAREI, R., HE, J., SIULY, S., AND ZHANG, Y. **A PCA aided cross-covariance scheme for discriminative feature extraction from EEG signals.** *Computer Methods and Programs in Biomedicine* 146 (2017), 47–57.
- [232] ZHANG, H., SUN, Y., LI, J., WANG, F., AND WANG, Z. **Covert Verb Reading Contributes to Signal Classification of Motor Imagery in BCI.** *IEEE Transactions on Neural Systems and Rehabilitation Engineering* 26, 1 (2018), 45–50.
- [233] AYDIN, S., ARICA, N., ERGUL, E., AND TAN, O. **Classification of Obsessive Compulsive Disorder by EEG Complexity and Hemispheric Dependency Measurements.** *International Journal of Neural Systems* 25, 03 (2015), 1550010.
- [234] BAE, Y., YOO, B. W., LEE, J. C., AND KIM, H. C. **Automated network analysis to measure brain effective connectivity estimated from EEG data of patients with alcoholism.** *Physiological Measurement* 38, 5 (2017), 759–773.
- [235] BAK, N., EBDROP, B. H., ORANJE, B., FAGERLUND, B., JENSEN, M. H., DÜRING, S. W., NIELSEN, M. O., GLENTHØJ, B. Y., AND HANSEN, L. K. **Two subgroups of antipsychotic-naive, first-episode schizophrenia patients identified with a Gaussian mixture model on cognition and electrophysiology.** *Translational Psychiatry* 7, 4 (2017), e1087–e1087.
- [236] BETROUNI, N., DELVAL, A., CHATON, L., DEFEBVRE, L., DUIJS, A., MOONEN, A., LEENTJENS, A. F., AND DUJARDIN, K. **Electroencephalography-based machine learning for cognitive profiling in Parkinson’s disease: Preliminary results.** *Movement Disorders* 34, 2 (2019), 210–217.
- [237] CAI, H., CHEN, Y., HAN, J., ZHANG, X., AND HU, B. **Study on Feature Selection Methods for Depression Detection Using Three-Electrode EEG Data.** *Interdisciplinary Sciences: Computational Life Sciences* 10, 3 (2018), 558–565.
- [238] CAO, Z., LAI, K.-L., LIN, C.-T., CHUANG, C.-H., CHOU, C.-C., AND WANG, S.-J. **Exploring resting-state EEG complexity before migraine attacks.** *Cephalalgia* 38, 7 (2018), 1296–1306.
- [239] CASTELHANO, J. A., TAVARES, P., MOUGA, S., OLIVEIRA, G., AND CASTELO-BRANCO, M. **Stimulus dependent neural oscillatory patterns show reliable statistical identification of autism spectrum disorder in a face perceptual decision task.** *Clinical Neurophysiology* 129, 5 (2018), 981–989.
- [240] DING, X., YUE, X., ZHENG, R., BI, C., LI, D., AND YAO, G. **Classifying major depression patients and healthy controls using EEG, eye tracking and galvanic skin response data.** *Journal of Affective Disorders* 251 (2019), 156–161.

- [241] GRAM, M., GRAVERSEN, C., OLESEN, A., AND DREWES, A. **Machine learning on encephalographic activity may predict opioid analgesia.** *European Journal of Pain* 19, 10 (2015), 1552–1561.
- [242] HÜLLER, Y., BATHKE, A. C., UHL, A., STROBL, N., LANG, A., BERGMANN, J., NARDONE, R., ROSSINI, F., ZAUNER, H., KIRSCHNER, M., JAHANBEKAM, A., TRINKA, E., AND STAFFEN, W. **Combining SPECT and Quantitative EEG Analysis for the Automated Differential Diagnosis of Disorders with Amnestic Symptoms.** *Frontiers in Aging Neuroscience* 9 (2017), 290.
- [243] HOUMANI, N., VIALATTE, F., GALLEGRO-JUTGÉ, E., DREYFUS, G., NGUYEN-MICHEL, V.-H., MARIANI, J., AND KINUGAWA, K. **Diagnosis of Alzheimer’s disease with Electroencephalography in a differential framework.** *PLOS ONE* 13, 3 (2018), e0193607.
- [244] IERACITANO, C., MAMMONE, N., HUSSAIN, A., AND MORABITO, F. C. **A novel multi-modal machine learning based approach for automatic classification of EEG recordings in dementia.** *Neural Networks* 123 (2020), 176–190.
- [245] JAHMUNAH, V., LIH OH, S., RAJINIKANTH, V., CIACCIO, E. J., HAO CHEONG, K., ARUNKUMAR, N., AND ACHARYA, U. R. **Automated detection of schizophrenia using nonlinear signal processing methods.** *Artificial Intelligence in Medicine* 100 (2019), 101698.
- [246] KALATZIS, I., PILIOURAS, N., VENTOURAS, E., PAPAGEORGIOU, C., RABAVILAS, A., AND CAVOURAS, D. **Design and implementation of an SVM-based computer classification system for discriminating depressive patients from healthy controls using the P600 component of ERP signals.** *Computer Methods and Programs in Biomedicine* 75, 1 (2004), 11–22.
- [247] KHATUN, S., MORSHED, B. I., AND BIDELMAN, G. M. **A Single-Channel EEG-Based Approach to Detect Mild Cognitive Impairment via Speech-Evoked Brain Responses.** *IEEE Transactions on Neural Systems and Rehabilitation Engineering* 27, 5 (2019), 1063–1070.
- [248] KNOTH, I. S., LAJNEF, T., RIGOULOT, S., LACOURSE, K., VANNASING, P., MICHAUD, J. L., JACQUEMONT, S., MAJOR, P., JERBI, K., AND LIPPÉ, S. **Auditory repetition suppression alterations in relation to cognitive functioning in fragile X syndrome: a combined EEG and machine learning approach.** *Journal of Neurodevelopmental Disorders* 10, 1 (2018), 4.
- [249] LI, X., HU, B., SHEN, J., XU, T., AND RETCLIFFE, M. **Mild Depression Detection of College Students: an EEG-Based Solution with Free Viewing Tasks.** *Journal of Medical Systems* 39, 12 (2015), 187.
- [250] LI, X., HU, B., SUN, S., AND CAI, H. **EEG-based mild depressive detection using feature selection methods and classifiers.** *Computer Methods and Programs in Biomedicine* 136 (2016), 151–161.
- [251] LIAO, S.-C., WU, C.-T., HUANG, H.-C., CHENG, W.-T., AND LIU, Y.-H. **Major Depression Detection from EEG Signals Using Kernel Eigen-Filter-Bank Common Spatial Patterns.** *Sensors* 17, 6 (2017), 1385.

- [252] MAHATO, S., AND PAUL, S. **Classification of Depression Patients and Normal Subjects Based on Electroencephalogram (EEG) Signal Using Alpha Power and Theta Asymmetry.** *Journal of Medical Systems* 44, 1 (2020), 28.
- [253] MOHAGHEGHIAN, F. **Computer-Aided Tinnitus Detection based on Brain Network Analysis of EEG Functional Connectivity.** *Journal of Biomedical Physics and Engineering* 9, 6 (2019).
- [254] MUMTAZ, W., ALI, S. S. A., YASIN, M. A. M., AND MALIK, A. S. **A machine learning framework involving EEG-based functional connectivity to diagnose major depressive disorder (MDD).** *Medical and Biological Engineering and Computing* 56, 2 (2018), 233–246.
- [255] MUMTAZ, W., SAAD, M. N. B. M., KAMEL, N., ALI, S. S. A., AND MALIK, A. S. **An EEG-based functional connectivity measure for automatic detection of alcohol use disorder.** *Artificial Intelligence in Medicine* 84 (2018), 79–89.
- [256] PIRYATINSKA, A., DARKHOVSKY, B., AND KAPLAN, A. **Binary classification of multichannel-EEG records based on the ϵ -complexity of continuous vector functions.** *Computer Methods and Programs in Biomedicine* 152 (2017), 131–139.
- [257] SÁNCHEZ-GONZÁLEZ, A., GARCÍA-ZAPIRAIN, B. N., MAESTRO SAIZ, I., AND YURREBASO SANTAMARÍA, I. **Patient prognosis based on feature extraction, selection and classification of EEG periodic activity.** *Bio-Medical Materials and Engineering* 26, s1 (2015), S1569–S1578.
- [258] STRAUSS, D., DELB, W., AND PLINKERT, P. **Objective Detection of the Central Auditory Processing Disorder: A New Machine Learning Approach.** *IEEE Transactions on Biomedical Engineering* 51, 7 (2004), 1147–1155.
- [259] TEKIN ERGUZEL, T., TAS, C., AND CEBI, M. **A wrapper-based approach for feature selection and classification of major depressive disorder-bipolar disorders.** *Computers in Biology and Medicine* 64 (2015), 127–137.
- [260] TORABI, A., DALIRI, M. R., AND SABZPOSHAN, S. H. **Diagnosis of multiple sclerosis from EEG signals using nonlinear methods.** *Australasian Physical & Engineering Sciences in Medicine* 40, 4 (2017), 785–797.
- [261] TRAMBAIOLLI, L., SPOLAÔR, N., LORENA, A., ANGHINAH, R., AND SATO, J. **Feature selection before EEG classification supports the diagnosis of Alzheimer’s disease.** *Clinical Neurophysiology* 128, 10 (2017), 2058–2067.
- [262] YU, H., LI, X., LEI, X., AND WANG, J. **Modulation Effect of Acupuncture on Functional Brain Networks and Classification of Its Manipulation With EEG Signals.** *IEEE Transactions on Neural Systems and Rehabilitation Engineering* 27, 10 (2019), 1973–1984.
- [263] ZARE, M., REZVANI, Z., AND BENASICH, A. A. **Automatic classification of 6-month-old infants at familial risk for language-based learning disorder using a support vector machine.** *Clinical Neurophysiology* 127, 7 (2016), 2695–2703.

- [264] IBRAHIM, I. A., TING, H.-N., AND MOGHAVVEMI, M. **A Novel Approach for Classifying Native Chinese and Malay Speaking Persons According to Cortical Auditory Evoked Responses.** *The Journal of International Advanced Otolaryngology* 15, 1 (2019), 87–93.
- [265] JAWED, S., AMIN, H. U., MALIK, A. S., AND FAYE, I. **Classification of Visual and Non-visual Learners Using Electroencephalographic Alpha and Gamma Activities.** *Frontiers in Behavioral Neuroscience* 13 (2019), 86.
- [266] SAI, C. Y., MOKHTAR, N., AROF, H., CUMMING, P., AND IWAHASHI, M. **Automated Classification and Removal of EEG Artifacts With SVM and Wavelet-ICA.** *IEEE Journal of Biomedical and Health Informatics* 22, 3 (2018), 664–670.
- [267] URIGÜEN, J. A., AND GARCIA-ZAPIRAIN, B. **EEG artifact removal—state-of-the-art and guidelines.** *Journal of Neural Engineering* 12, 3 (2015), 031001.
- [268] YANG, B., DUAN, K., AND ZHANG, T. **Removal of EOG artifacts from EEG using a cascade of sparse autoencoder and recursive least squares adaptive filter.** *Neurocomputing* 214 (2016), 1053–1060.
- [269] NOLAN, H., WHELAN, R., AND REILLY, R. **FASTER: Fully automated statistical thresholding for EEG artifact rejection.** *Journal of Neuroscience Methods* 192, 1 (2010), 152–162.
- [270] DEROUESNÉ, C., AND LACOMBLEZ, L. **Depression and dementia.** *Psychologie & neuropsychiatrie du vieillissement* 2 Suppl 1 (2004), S35–S42.
- [271] PLATEL, H., AND GROUSSARD, M. **La mémoire sémantique musicale : apport des données de la neuropsychologie clinique et de la neuro-imagerie fonctionnelle.** *Revue de neuropsychologie* 2, 1 (2010), 61.
- [272] CRYSTAL, H. A., GROBER, E., AND MASUR, D. **Preservation of musical memory in Alzheimer's disease.** *Journal of Neurology, Neurosurgery & Psychiatry* 52, 12 (1989), 1415–1416.
- [273] CARLESIMO, G. A., AND OSCAR-BERMAN, M. **Memory deficits in Alzheimer's patients: A comprehensive review.** *Neuropsychology Review* 3, 2 (1992), 119–169.
- [274] POLK, M., AND KERTESZ, A. **Music and language in degenerative disease of the brain.** *Brain and Cognition* 22, 1 (1993), 98–117.
- [275] COWLES, A., BEATTY, W. W., NIXON, S. J., LUTZ, L. J., PAULK, J., PAULK, K., AND ROSS, E. D. **Musical Skill in Dementia: A Violinist Presumed to Have Alzheimer's Disease Learns to Play a New Song.** *Neurocase* 9, 6 (2003), 493–503.
- [276] CUDDY, L. L., AND DUFFIN, J. **Music, memory, and Alzheimer's disease: is music recognition spared in dementia, and how can it be assessed?** *Medical Hypotheses* 64, 2 (2005), 229–235.
- [277] MOUSSARD, A., BIGAND, E., CLÉMENT, S., AND SAMSON, S. **Préservation des apprentissages implicites en musique dans le vieillissement normal et la maladie d'Alzheimer.** *Revue De Neuropsychologie* 18 (2008), 127–152.

- [278] BARTLETT, J. C., HALPERN, A. R., AND DOWLING, W. J. **Recognition of familiar and unfamiliar melodies in normal aging and Alzheimer's disease.** *Memory & Cognition* 23, 5 (1995), 531–546.
- [279] GROUSSARD, M., VIADER, F., HUBERT, V., LANDEAU, B., ABBAS, A., DESGRANGES, B., EUSTACHE, F., AND PLATEL, H. **Musical and verbal semantic memory: Two distinct neural networks?** *NeuroImage* 49, 3 (2010), 2764–2773.
- [280] JACOBSEN, J.-H., STELZER, J., FRITZ, T. H., CHÉTELAT, G., JOIE, R. L., AND TURNER, R. **Why musical memory can be preserved in advanced Alzheimer's disease.** *Brain* 138, 8 (2015), 2438–2450.
- [281] SAMSON, S., DELLACHERIE, D., AND PLATEL, H. **Emotional power of music in patients with memory disorders.** *Annals of the New York Academy of Sciences* 1169, 1 (2009), 245–255.
- [282] WITTEKIND, C. E., TERFEHR, K., OTTE, C., JELINEK, L., HINKELMANN, K., AND MORITZ, S. **Mood-congruent memory in depression – the influence of personal relevance and emotional context.** *Psychiatry Research* 215, 3 (2014), 606–613.
- [283] CLEEREMANS, A., DESTREBECQZ, A., AND BOYER, M. **Implicit learning: news from the front.** *Trends in Cognitive Sciences* 2, 10 (1998), 406–416.
- [284] KRAEMER, D. J. M., MACRAE, C. N., GREEN, A. E., AND KELLEY, W. M. **Sound of silence activates auditory cortex.** *Nature* 434, 7030 (2005), 158–158.
- [285] HALPERN, A. R. **Mental scanning in auditory imagery for songs.** *Journal of Experimental Psychology: Learning, Memory, and Cognition* 14, 3 (1988), 434–443.
- [286] POVEL, D.-J., AND ESSENS, P. **Perception of temporal patterns.** *Music Perception* 2, 4 (1985), 411–440.
- [287] LEVITIN, D. J., AND COOK, P. R. **Memory for musical tempo: Additional evidence that auditory memory is absolute.** *Perception & Psychophysics* 58, 6 (1996), 927–935.
- [288] HALPERN, A. R. **Memory for the absolute pitch of familiar songs.** *Memory & Cognition* 17, 5 (1989), 572–581.
- [289] JANATA, P., AND PAROO, K. **Acuity of auditory images in pitch and time.** *Perception and Psychophysics* 68, 5 (2006), 829–844.
- [290] HALPERN, A. R., ZATORRE, R. J., BOUFFARD, M., AND JOHNSON, J. A. **Behavioral and neural correlates of perceived and imagined musical timbre.** *Neuropsychologia* 42, 9 (2004), 1281–1292.
- [291] LUCAS, B. J., SCHUBERT, E., AND HALPERN, A. R. **Perception of emotion in sounded and imagined music.** *Music Perception* 27, 5 (2010), 399–412.
- [292] GABRIEL, D., WONG, T. C., NICOLIER, M., GIUSTINIANI, J., MIGNOT, C., NOIRET, N., MONNIN, J., MAGNIN, E., PAZART, L., MOULIN, T., HAFFEN, E., AND VANDEL, P. **Don't forget the lyrics! Spatiotemporal dynamics of neural mechanisms spontaneously evoked by gaps of silence in familiar and newly learned songs.** *Neurobiology of Learning and Memory* 132 (2016), 18–28.

- [293] SCHÖNWIESNER, M., NOVITSKI, N., PAKARINEN, S., CARLSON, S., TERVANIEMI, M., AND NÄÄTÄNEN, R. **Heschl's gyrus, posterior superior temporal gyrus, and mid-ventrolateral prefrontal cortex have different roles in the detection of acoustic changes.** *Journal of Neurophysiology* 97, 3 (2007), 2075–2082.
- [294] BOUWER, F. L., ZUIJEN, T. L. V., AND HONING, H. **Beat processing is pre-attentive for metrically simple rhythms with clear accents: An ERP study.** *PLoS ONE* 9, 5 (2014), e97467.
- [295] FOLSTEIN, J. R., AND PETTEN, C. V. **Influence of cognitive control and mismatch on the N2 component of the ERP: A review.** *Psychophysiology* 45, 1 (2007), 152–170.
- [296] HUANG, M.-X., LEE, R. R., GAA, K. M., SONG, T., HARRINGTON, D. L., LOH, C., THEILMANN, R. J., EDGAR, J. C., MILLER, G. A., CANIVE, J. M., AND GRANHOLM, E. **Somatosensory system deficits in schizophrenia revealed by MEG during a median-nerve oddball task.** *Brain Topography* 23, 1 (2009), 82–104.
- [297] PEREIRA, C. S., TEIXEIRA, J., FIGUEIREDO, P., XAVIER, J., CASTRO, S. L., AND BRATTICO, E. **Music and emotions in the brain: Familiarity matters.** *PLoS ONE* 6, 11 (2011), e27241.
- [298] HALPERN, A. R. **When that tune runs through your head: A PET investigation of auditory imagery for familiar melodies.** *Cerebral Cortex* 9, 7 (1999), 697–704.
- [299] JANATA, P. **ERP measures assay the degree of expectancy violation of harmonic contexts in music.** *Journal of Cognitive Neuroscience* 7, 2 (1995), 153–164.
- [300] FRIEDMAN, D., AND JOHNSON, R. **Event-related potential (ERP) studies of memory encoding and retrieval: A selective review.** *Microscopy Research and Technique* 51, 1 (2000), 6–28.
- [301] JONGSMA, M. L., GERRITS, N. J., VAN RIJN, C. M., QUIROGA, R. Q., AND MAES, J. H. **Event related potentials to digit learning: Tracking neurophysiologic changes accompanying recall performance.** *International Journal of Psychophysiology* 85, 1 (2012), 41–48.
- [302] POLICH, J. **Updating P300: An integrative theory of P3a and P3b.** *Clinical Neurophysiology* 118, 10 (2007), 2128–2148.
- [303] KELLENBACH, M. L., WIJERS, A. A., AND MULDER, G. **Visual semantic features are activated during the processing of concrete words: event-related potential evidence for perceptual semantic priming.** *Cognitive Brain Research* 10, 1-2 (2000), 67–75.
- [304] ZHANG, Q., LAWSON, A., GUO, C., AND JIANG, Y. **Electrophysiological correlates of visual affective priming.** *Brain Research Bulletin* 71, 1-3 (2006), 316–323.
- [305] HURON, D. **Sweet Anticipation: Music and the Psychology of Expectation.** The MIT Press, 2006.
- [306] SCHÄFER, T., AND SEDLMEIER, P. **From the functions of music to music preference.** *Psychology of Music* 37, 3 (2009), 279–300.

- [307] RENTFROW, P. J., AND GOSLING, S. D. **The do re mi's of everyday life: The structure and personality correlates of music preferences.** *Journal of Personality and Social Psychology* 84, 6 (2003), 1236–1256.
- [308] SALIMPOOR, V. N., ZALD, D. H., ZATORRE, R. J., DAGHER, A., AND MCINTOSH, A. R. **Predictions and the brain: how musical sounds become rewarding.** *Trends in Cognitive Sciences* 19, 2 (2015), 86–91.
- [309] SCHÄFER, T., ZIMMERMANN, D., AND SEDLMEIER, P. **How we remember the emotional intensity of past musical experiences.** *Frontiers in Psychology* 5 (2014).
- [310] SCHUBERT, E., HARGREAVES, D. J., AND NORTH, A. C. **A dynamically minimalist cognitive explanation of musical preference: is familiarity everything?** *Frontiers in Psychology* 5 (2014).
- [311] VUUST, P., OSTERGAARD, L., PALLESEN, K. J., BAILEY, C., AND ROEPSTORFF, A. **Predictive coding of music – brain responses to rhythmic incongruity.** *Cortex* 45, 1 (2009), 80–92.
- [312] VUUST, P., AND WITEK, M. A. G. **Rhythmic complexity and predictive coding: a novel approach to modeling rhythm and meter perception in music.** *Frontiers in Psychology* 5 (2014).
- [313] SCHAEFER, R. S., VLEK, R. J., AND DESAIN, P. **Music perception and imagery in EEG: Alpha band effects of task and stimulus.** *International Journal of Psychophysiology* 82, 3 (2011), 254–259.
- [314] VLEK, R., SCHAEFER, R., GIELEN, C., FARQUHAR, J., AND DESAIN, P. **Shared mechanisms in perception and imagery of auditory accents.** *Clinical Neurophysiology* 122, 8 (2011), 1526–1532.
- [315] ZATORRE, R. J., HALPERN, A. R., PERRY, D. W., MEYER, E., AND EVANS, A. C. **Hearing in the mind's ear: A PET investigation of musical imagery and perception.** *Journal of Cognitive Neuroscience* 8, 1 (1996), 29–46.
- [316] GUITRY, S. **Toutes rélexions faites.** In *Cinquante ans d'occupations*, Omnibus. Presses de la cité, 1947.
- [317] LINKE, A. C., AND CUSACK, R. **Flexible information coding in human auditory cortex during perception, imagery, and STM of complex sounds.** *Journal of Cognitive Neuroscience* 27, 7 (2015), 1322–1333.
- [318] BLOOD, A. J., AND ZATORRE, R. J. **Intensely pleasurable responses to music correlate with activity in brain regions implicated in reward and emotion.** *Proceedings of the National Academy of Sciences* 98, 20 (2001), 11818–11823.
- [319] STALINSKI, S. M., AND SCHELLENBERG, E. G. **Listeners remember music they like.** *Journal of Experimental Psychology: Learning, Memory, and Cognition* 39, 3 (2013), 700–716.
- [320] SALIMPOOR, V. N., BENOVOY, M., LARCHER, K., DAGHER, A., AND ZATORRE, R. J. **Anatomically distinct dopamine release during anticipation and experience of peak emotion to music.** *Nature Neuroscience* 14, 2 (2011), 257–262.

- [321] CLARK, C. N., DOWNEY, L. E., AND WARREN, J. D. **Music biology: All this useful beauty.** *Current Biology* 24, 6 (2014), R234–R237.
- [322] KOELSCH, S., FRITZ, T., V. CRAMON, D. Y., MÜLLER, K., AND FRIEDERICI, A. D. **Investigating emotion with music: An fMRI study.** *Human Brain Mapping* 27, 3 (2006), 239–250.
- [323] SAMMLER, D., GRIGUTSCH, M., FRITZ, T., AND KOELSCH, S. **Music and emotion: Electrophysiological correlates of the processing of pleasant and unpleasant music.** *Psychophysiology* 44, 2 (2007), 293–304.
- [324] HADJIDIMITRIOU, S. K., AND HADJILEONTIADIS, L. J. **Toward an EEG-based recognition of music liking using time-frequency analysis.** *IEEE Transactions on Biomedical Engineering* 59, 12 (2012), 3498–3510.
- [325] HADJIDIMITRIOU, S. K., AND HADJILEONTIADIS, L. J. **EEG-based classification of music appraisal responses using time-frequency analysis and familiarity ratings.** *IEEE Transactions on Affective Computing* 4, 2 (2013), 161–172.
- [326] ADAMOS, D. A., DIMITRIADIS, S. I., AND LASKARIS, N. A. **Towards the bio-personalization of music recommendation systems: A single-sensor EEG biomarker of subjective music preference.** *Information Sciences* 343-344 (2016), 94–108.
- [327] DELORME, A., AND MAKEIG, S. **EEGLAB: an open source toolbox for analysis of single-trial EEG dynamics including independent component analysis.** *Journal of Neuroscience Methods* 134, 1 (2004), 9–21.
- [328] DE PERALTA MENENDEZ, R. G., ANDINO, S. G., LANTZ, G., MICHEL, C. M., AND LANDIS, T. **Noninvasive localization of electromagnetic epileptic activity, I. Method descriptions and simulations.** *Brain Topography* 14, 2 (2001), 131–137.
- [329] CROWDER, R. G., SERAFINE, M. L., AND REPP, B. **Physical interaction and association by contiguity in memory for the words and melodies of songs.** *Memory & Cognition* 18, 5 (1990), 469–476.
- [330] PERETZ, I., AND ZATORRE, R. J. **Brain organization for music processing.** *Annual Review of Psychology* 56, 1 (2005), 89–114.
- [331] HUNTER, P. G., SCHELLENBERG, E. G., AND SCHIMMACK, U. **Mixed affective responses to music with conflicting cues.** *Cognition & Emotion* 22, 2 (2008), 327–352.
- [332] VUILLEUMIER, P., AND TROST, W. **Music and emotions: from enchantment to entrainment.** *Annals of the New York Academy of Sciences* 1337, 1 (2015), 212–222.
- [333] BRATTICO, E., BOGERT, B., ALLURI, V., TERVANIEMI, M., EEROLA, T., AND JACOBSEN, T. **It's sad but i like it: The neural dissociation between musical emotions and liking in experts and laypersons.** *Frontiers in Human Neuroscience* 9 (2016).
- [334] BURGESS, A. P., AND GRUZELIER, J. H. **Short duration synchronization of human theta rhythm during recognition memory.** *NeuroReport* 8, 4 (1997), 1039–1042.

- [335] INANAGA, K. **Frontal midline theta rhythm and mental activity.** *Psychiatry and Clinical Neurosciences* 52, 6 (1998), 555–566.
- [336] SMITH, M. E., MCEVOY, L. K., AND GEVINS, A. **Neurophysiological indices of strategy development and skill acquisition.** *Cognitive Brain Research* 7, 3 (1999), 389–404.
- [337] LANG, P. J., BRADLEY, M. M., AND CUTHBERT, B. N. **Motivated attention: Affect, activation, and action.** In *Attention and Orienting: Sensory and Motivational Processes*, P. J. Lang, R. F. Simons, and M. T. Balaban, Eds. Lawrence Erlbaum Associates, Hillsdale, NJ, 1997, pp. 97–135.
- [338] KLIMESCH, W., SAUSENG, P., HANSLMAYR, S., GRUBER, W., AND FREUNBERGER, R. **Event-related phase reorganization may explain evoked neural dynamics.** *Neuroscience & Biobehavioral Reviews* 31, 7 (2007), 1003–1016.
- [339] CALDWELL, G. N., AND RIBY, L. M. **The effects of music exposure and own genre preference on conscious and unconscious cognitive processes: A pilot ERP study.** *Consciousness and Cognition* 16, 4 (2007), 992–996.
- [340] FUJIOKA, T., TRAINOR, L. J., LARGE, E. W., AND ROSS, B. **Beta and gamma rhythms in human auditory cortex during musical beat processing.** *Annals of the New York Academy of Sciences* 1169, 1 (2009), 89–92.
- [341] EDAGAWA, K., AND KAWASAKI, M. **Beta phase synchronization in the frontal-temporal-cerebellar network during auditory-to-motor rhythm learning.** *Scientific Reports* 7, 1 (2017).
- [342] GORDON, R. L., MAGNE, C. L., AND LARGE, E. W. **EEG correlates of song prosody: A new look at the relationship between linguistic and musical rhythm.** *Frontiers in Psychology* 2 (2011).
- [343] PLATEL, H., BARON, J.-C., DESGRANGES, B., BERNARD, F., AND EUSTACHE, F. **Semantic and episodic memory of music are subserved by distinct neural networks.** *NeuroImage* 20, 1 (2003), 244–256.
- [344] SATOH, M., TAKEDA, K., NAGATA, K., SHIMOSEGAWA, E., AND KUZUHARA, S. **Positron-emission tomography of brain regions activated by recognition of familiar music.** *AJNR. American journal of neuroradiology* 27 (2006), 1101–1106.
- [345] ALTENMÜLLER, E., SIGGEL, S., MOHAMMADI, B., SAMII, A., AND MÜNTE, T. F. **Play it again sam: brain correlates of emotional music recognition.** *Frontiers in Psychology* 5 (2014).
- [346] WILKINS, R. W., HODGES, D. A., LAURIENTI, P. J., STEEN, M., AND BURDETTE, J. H. **Network Science and the Effects of Music Preference on Functional Brain Connectivity: From Beethoven to Eminem.** *Scientific Reports* 4, 1 (2014).
- [347] SWANN, N., TANDON, N., CANOLTY, R., ELLMORE, T. M., MCEVOY, L. K., DREYER, S., DISANO, M., AND ARON, A. R. **Intracranial EEG reveals a time- and frequency-specific role for the right inferior frontal gyrus and primary motor cortex in stopping initiated responses.** *Journal of Neuroscience* 29, 40 (2009), 12675–12685.

- [348] SUZUKI, A. **Emotional functions of the insula.** *Brain Nerve* 64 (2012), 1103–1112.
- [349] SCHELLENBERG, E. G., PERETZ, I., AND VIEILLARD, S. **Liking for happy- and sad-sounding music: Effects of exposure.** *Cognition & Emotion* 22, 2 (2008), 218–237.
- [350] SZPUNAR, K. K., SCHELLENBERG, E. G., AND PLINER, P. **Liking and memory for musical stimuli as a function of exposure.** *Journal of Experimental Psychology: Learning, Memory, and Cognition* 30, 2 (2004), 370–381.
- [351] SHALGI, S., AND DEQUELL, L. Y. **Direct evidence for differential roles of temporal and frontal components of auditory change detection.** *Neuropsychologia* 45, 8 (2007), 1878–1888.
- [352] VAN GERVEN, M., FARQUHAR, J., SCHAEFER, R., VLEK, R., GEUZE, J., NIJHOLT, A., RAMSEY, N., HASELAGER, P., VUURPIJL, L., GIELEN, S., AND DESAIN, P. **The brain–computer interface cycle.** *Journal of neural engineering* 6 (2009), 041001.
- [353] MÜLLER-PUTZ, G. R., SCHERER, R., BRUNNER, C., LEEB, R., AND PFURTSCHELLER, G. **Better than random: a closer look on BCI results.** *International Journal of Bioelectromagnetism* (2008).
- [354] HUA, J., XIONG, Z., LOWEY, J., SUH, E., AND DOUGHERTY, E. R. **Optimal number of features as a function of sample size for various classification rules.** *Bioinformatics* 21, 8 (2004), 1509–1515.
- [355] HUA, J., XIONG, Z., AND DOUGHERTY, E. R. **Determination of the optimal number of features for quadratic discriminant analysis via the normal approximation to the discriminant distribution.** *Pattern Recognition* 38, 3 (2005), 403–421.
- [356] HERHOLZ, S. C., AND ZATORRE, R. J. **Musical training as a framework for brain plasticity: Behavior, function, and structure.** *Neuron* 76, 3 (2012), 486 – 502.
- [357] HÖLLER, Y., BERGMANN, J., THOMSCHEWSKI, A., KRONBICHLER, M., HÖLLER, P., CRONE, J. S., SCHMID, E. V., BUTZ, K., NARDONE, R., AND TRINKA, E. **Comparison of EEG-features and classification methods for motor imagery in patients with disorders of consciousness.** *PLOS ONE* 8, 11 (2013), 1–15.
- [358] WAINER, J., AND FONSECA, P. **How to tune the RBF SVM hyperparameters?: An empirical evaluation of 18 search algorithms.** *arXiv*, 2008.11655v1 (2020).
- [359] ROEST, A., VAN BETS, B., JORENS, P. G., BAAR, I., WEYLER, J., AND MERCELIS, R. **The prognostic value of the EEG in postanoxic coma.** *Neurocritical Care* 10, 3 (2009), 318–325.
- [360] LOGI, F., FISCHER, C., MURRI, L., AND MAUGUIÈRE, F. **The prognostic value of evoked responses from primary somatosensory and auditory cortex in comatose patients.** *Clinical Neurophysiology* 114, 9 (2003), 1615–1627.
- [361] KOTCHOUBEY, B., LANG, S., MEZGER, G., SCHMALOHR, D., SCHNECK, M., SEMMLER, A., BOSTANOV, V., AND BIRBAUMER, N. **Information processing in severe disorders of consciousness: Vegetative state and minimally conscious state.** *Clinical Neurophysiology* 116, 10 (2005), 2441–2453.

- [362] HINTERBERGER, T., WILHELM, B., MELLINGER, J., KOTCHOUBEY, B., AND BIRBAUMER, N. **A device for the detection of cognitive brain functions in completely paralyzed or unresponsive patients.** *IEEE Transactions on Biomedical Engineering* 52, 2 (2005), 211–220.
- [363] NEUMANN, N., AND KOTCHOUBEY, B. **Assessment of cognitive functions in severely paralysed and severely brain-damaged patients: neuropsychological and electrophysiological methods.** *Brain Research Protocols* 14, 1 (2004), 25–36.
- [364] GLASS, I., SAZBON, L., AND GROSWASSER, Z. **Mapping “cognitive” event-related potentials in prolonged postcoma unawareness state.** *Clinical Electroencephalography* 29, 1 (1998), 19–30.
- [365] MUTSCHLER, V., CHAUMEIL, C. G., MARCOUX, L., WIOLAND, N., TEMPÉ, J., AND KURTZ, D. **Étude du P300 auditif chez des sujets en coma post-anoxique. Données préliminaires.** *Neurophysiologie Clinique/Clinical Neurophysiology* 26, 3 (1996), 158–163.
- [366] FISCHER, C., LUAUTE, J., SALORD, F., JOURDAN, C., AND MORLET, D. **Valeur pronostique des PEA au stade aigu du coma. PE auditifs précoces, de latence moyenne, tardifs et négativité de discordance (MMN).** In *L'évaluation neurophysiologique des comas, de la mort encéphalique et des états végétatifs*, J. Guerit, Ed. Solal, Marseille, 2001, pp. 169–181.
- [367] FISCHER, C., LUAUTE, J., ADELEINE, P., AND MORLET, D. **Predictive value of sensory and cognitive evoked potentials for awakening from coma.** *Neurology* 63, 4 (2004), 669–673.
- [368] GUÉRIT, J. M., FISCHER, C., FACCO, E., TINUPER, P., MURRI, L., RONNE-ENGSTRÖM, E., AND NUWER, M. **Standards of clinical practice of EEG and EPs in comatose and other unresponsive states. The International Federation of Clinical Neurophysiology.** *Electroencephalography and clinical neurophysiology. Supplement* 52 (1999), 117–131.
- [369] KANE, N., CURRY, S., BUTLER, S., AND CUMMINS, B. **Electrophysiological indicator of awakening from coma.** *The Lancet* 341, 8846 (1993), 688.
- [370] KANE, N. M., CURRY, S. H., ROWLANDS, C. A., MANARA, A. R., LEWIS, T., MOSS, T., CUMMINS, B. H., AND BUTLER, S. R. **Event-related potentials — neurophysiological tools for predicting emergence and early outcome from traumatic coma.** *Intensive Care Medicine* 22, 1 (1996), 39–46.
- [371] SUPPIEJ, A., CAPPELLARI, A., AND COGO, P. **Prognostic role of somatosensory and auditory evoked potentials in paediatric hypoxic-ischemic encephalopathy managed with hypothermia: An illustrative case.** *Neurophysiologie Clinique/Clinical Neurophysiology* 39, 2 (2009), 101–105.
- [372] LUAUTÉ, J., FISCHER, C., ADELEINE, P., MORLET, D., TELL, L., AND BOISSON, D. **Late auditory and event-related potentials can be useful to predict good functional outcome after coma.** *Archives of Physical Medicine and Rehabilitation* 86, 5 (2005), 917–923.

- [373] KOTCHOUBEY, B., LANG, S., HERB, E., MAURER, P., SCHMALOHR, D., BOSTANOV, V., AND BIRBAUMER, N. **Stimulus complexity enhances auditory discrimination in patients with extremely severe brain injuries.** *Neuroscience Letters* 352, 2 (2003), 129–132.
- [374] KOTCHOUBEY, B. **Event-related potential measures of consciousness: two equations with three unknowns.** In *Progress in Brain Research*. Elsevier, 2005, pp. 427–444.
- [375] WIJNEN, V., VAN BOXTEL, G., EILANDER, H., AND DE GELDER, B. **Mismatch negativity predicts recovery from the vegetative state.** *Clinical Neurophysiology* 118, 3 (2007), 597–605.
- [376] SCHNAKERS, C., PERRIN, F., SCHABUS, M., MAJERUS, S., LEDOUX, D., DAMAS, P., BOLY, M., VANHAUDENHUYSE, A., BRUNO, M. A., MOONEN, G., AND LAUREYS, S. **Voluntary brain processing in disorders of consciousness.** *Neurology* 71, 20 (2008), 1614–1620.
- [377] SCHOENLE, P. W., AND WITZKE, W. **How vegetative is the vegetative state? Preserved semantic processing in VS patients—evidence from N400 event-related potentials.** *NeuroRehabilitation* 19 (2004), 329–334.
- [378] DALTROZZO, J., WIOLAND, N., MUTSCHLER, V., LUTUN, P., CALON, B., MEYER, A., POTTECHER, T., LANG, S., JAEGER, A., AND KOTCHOUBEY, B. **Cortical information processing in coma.** *Cognitive and Behavioral Neurology* 22, 1 (2009), 53–62.
- [379] COHEN, M. X. **The discrete time Fourier transform, the FFT, and the convolution theorem.** In *Analyzing Neural Time Series Data: Theory and Practice*. MIT Press, 2014.
- [380] HINK, R., VOORHIS, S. V., HILLYARD, S., AND SMITH, T. **The division of attention and the human auditory evoked potential.** *Neuropsychologia* 15, 4-5 (1977), 597–605.
- [381] CROWLEY, K. E., AND COLRAIN, I. M. **A review of the evidence for P2 being an independent component process: age, sleep and modality.** *Clinical Neurophysiology* 115, 4 (2004), 732–744.
- [382] KUTAS, M., AND FEDERMEIER, K. D. **Thirty Years and Counting: Finding Meaning in the N400 Component of the Event-Related Brain Potential (ERP).** *Annual Review of Psychology* 62, 1 (2011), 621–647.
- [383] DEACON, D. **How automatically is meaning accessed: a review of the effects of attention on semantic processing.** *Frontiers in Bioscience* 5, 1 (2000), e82.
- [384] MCCARTHY, G., AND NOBRE, A. C. **Modulation of semantic processing by spatial selective attention.** *Electroencephalography and Clinical Neurophysiology/Evoked Potentials Section* 88, 3 (1993), 210–219.
- [385] STENBERG, G., LINDGREN, M., JOHANSSON, M., OLSSON, A., AND ROSÉN, I. **Semantic processing without conscious identification: Evidence from event-related potentials.** *Journal of Experimental Psychology: Learning, Memory, and Cognition* 26, 4 (2000), 973–1004.

- [386] JUOTTONEN, K., REVONSUO, A., AND LANG, H. **Dissimilar age influences on two ERP waveforms (LPC and N400) reflecting semantic context effect.** *Cognitive Brain Research* 4, 2 (1996), 99–107.
- [387] DUZEL, E., YONELINAS, A. P., MANGUN, G. R., HEINZE, H.-J., AND TULVING, E. **Event-related brain potential correlates of two states of conscious awareness in memory.** *Proceedings of the National Academy of Sciences* 94, 11 (1997), 5973–5978.
- [388] MISRA, M., AND HOLCOMB, P. J. **Event-related potential indices of masked repetition priming.** *Psychophysiology* 40, 1 (2003), 115–130.
- [389] DALTROZZO, J., WIOLAND, N., AND KOTCHOUBEY, B. **Sex differences in two event-related potentials components related to semantic priming.** *Archives of Sexual Behavior* 36, 4 (2007), 555–568.
- [390] CRISTIANINI, N., AND SHAWE-TAYLOR, J. **An Introduction to Support Vector Machines and Other Kernel-based Learning Methods.** Cambridge University Press, 2000.
- [391] CUCKER, F., AND SMALE, S. **On the mathematical foundations of learning.** *Bulletin of the American Mathematical Society* 39, 01 (2001), 1–50.
- [392] CAWLEY, G. C., AND TALBOT, N. L. C. **Preventing over-fitting during model selection via bayesian regularisation of the hyper-parameters.** *Journal of Machine Learning Research* 8 (2007), 841–861.
- [393] STONE, M. **Cross-validatory choice and assessment of statistical predictions.** *Journal of the Royal Statistical Society: Series B (Methodological)* 36, 2 (1974), 111–133.
- [394] FANELLI, D. **Opinion: Is science really facing a reproducibility crisis, and do we need it to?** *Proceedings of the National Academy of Sciences* 115, 11 (2018), 2628–2631.
- [395] BAKER, M. **1,500 scientists lift the lid on reproducibility.** *Nature* 533, 7604 (2016), 452–454.
- [396] FIDLER, F., AND WILCOX, J. **Reproducibility of Scientific Results.** In *The Stanford Encyclopedia of Philosophy*, E. N. Zalta, Ed., winter 2018 ed. Metaphysics Research Lab, Stanford University, 2018.
- [397] BOSER, BERNHARD, E., GUYON, ISABELLE, M., AND VAPNIK, VLADIMIR, N. **A training algorithm for optimal margin classifiers.** *COLT '92: Proceedings of the fifth annual workshop on Computational learning theory* (1992), 144–152.
- [398] JAEGER, H., AND HAAS, H. **Harnessing nonlinearity: Predicting chaotic systems and saving energy in wireless communication.** *Science* 304, 5667 (2004), 78–80.
- [399] BOZHKOVA, L., KOPRINKOVA-HRISTOVA, P., AND GEORGIEVA, P. **Learning to decode human emotions with Echo State Networks.** *Neural Networks* 78 (2015), 112–119.

- [400] IBÁÑEZ SORIA, D., SORIA-FRISCH, A., GARCIA-OJALVO, J., AND RUFFINI, G. **Characterization of the non-stationary nature of steady-state visual evoked potentials using echo state networks.** *PLOS ONE* 14, 7 (2019), 1–13.
- [401] DANIEL, E., MEINDERTSMA, T., ARAZI, A., DONNER, T., AND DINSTEIN, I. **The relationship between trial-by-trial variability and oscillations of cortical population activity.** *Scientific Reports* 9 (2019), 16901.
- [402] TRENADO, C., GONZÁLEZ-RAMÍREZ, A., LIZÁRRAGA-CORTÉS, V., PEDROARENA LEAL, N., MANJARREZ, E., AND RUGE, D. **The potential of trial-by-trial variabilities of ongoing-EEG, evoked potentials, event related potentials and fMRI as diagnostic markers for neuropsychiatric disorders.** *Frontiers in Neuroscience* 12 (2019), 850.
- [403] ZHENG, M., YANG, B., AND XIE, Y. **EEG classification across sessions and across subjects through transfer learning in motor imagery-based brain-machine interface system.** *Medical and Biological Engineering and Computing* 58 (2020).
- [404] PERRIN, F. **Études électrophysiologiques de la discrimination sémantique pendant le sommeil.** PhD thesis, 2001. Thèse de doctorat dirigée par Garcia-Larrea, Luis et Bastuji, Hélène Neurosciences Lyon 1 2001.
- [405] DE CAMP, N. V., KALINKA, G., AND BERGELER, J. **Light-cured polymer electrodes for non-invasive EEG recordings.** *Scientific Reports* 8 (2018), 14041.
- [406] NORMAN, R., PAL, I., STEWART, C., WALSLEBEN, J., AND RAPOPORT, D. **Inter-observer agreement among sleep scorers from different centers in a large dataset.** *Sleep* 23 (2000), 901–908.
- [407] DANKER-HOPFE, H., KUNZ, D., GRUBER, G., KLÖSCH, G., LORENZO, J., HIMANEN, S., KEMP, B., PENZEL, T., R^ÖSCHKE, J., DORN, H., SCHLÖGL, A., TRENKER, E., AND DORFFNER, G. **Interrater reliability between scorers from eight european sleep laboratories in subjects with different sleep disorders.** *Journal of Sleep Research* 13 (2004), 63–69.
- [408] MUTO, V., BERTHOMIER, C., SCHMIDT, C., VANDEWALLE, G., JASPAR, M., DEVELLERS, J., CHELLAPPA, S., MEYER, C., PHILLIPS, C., BERTHOMIER, P., PRADO, J., BENOIT, O., BRANDEWINDER, M., MATTOU, J., AND MAQUET, P. **Inter- and intra-expert variability in sleep scoring: Comparison between visual and automatic analysis.** *Sleep* 41, suppl 1 (2018), A121.
- [409] ABOU JAOUDE, M., SUN, H., PELLERIN, K. R., PAVLOVA, M., SARKIS, R. A., CASH, S. S., WESTOVER, M. B., AND LAM, A. D. **Expert-level automated sleep staging of long-term scalp electroencephalography recordings using deep learning.** *Sleep* (2020).
- [410] FREY, C., AND OSBORNE, M. **The future of employment: How susceptible are jobs to computerisation?** *Oxford Martin* 114 (2013).
- [411] SARWAR, S., DENT, A., FAUST, K., RICHER, M., DJURIC, U., OMMEREN, R., AND DIAMANDIS, P. **Physician perspectives on integration of artificial intelligence into diagnostic pathology.** *NPJ Digital Medicine* 2 (2019), 28.

- [412] WANG, D., KHOSLA, A., GARGEYA, R., IRSHAD, H., AND BECK, A. H. **Deep learning for identifying metastatic breast cancer.** *arXiv*, 1606.05718 (2016).
- [413] GENET, M., LHEUREUX, F., AND TRUCHOT, D. **La relation soignant–soigné dans la littérature académique : tendances générales, limites et perspectives futures.** *Annales Médico-psychologiques, revue psychiatrique* 176 (2018), 135–144.

V

ANNEXES

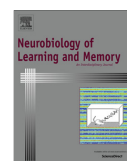
ANNEXE A : ARTICLE FONDATEUR DU PROJET RIHANNA

Neurobiology of Learning and Memory 132 (2016) 18–28



Contents lists available at ScienceDirect

Neurobiology of Learning and Memory

journal homepage: www.elsevier.com/locate/ynlme

Don't forget the lyrics! Spatiotemporal dynamics of neural mechanisms spontaneously evoked by gaps of silence in familiar and newly learned songs



Damien Gabriel^{a,b,*}, Thian Chiew Wong^a, Magali Nicolier^{a,b,c}, Julie Giustiniani^c, Coralie Mignot^a, Nicolas Noiret^{d,e}, Julie Monnin^{a,b,c}, Eloi Magnin^{d,f}, Lionel Pazart^{a,b}, Thierry Moulin^{a,b,f}, Emmanuel Haffen^{a,b,c}, Pierre Vandell^{a,b,c,d}

^a Centre d'investigation Clinique-Innovation Technologique CIC-IT 1431, Inserm, CHRU Besançon, F-25000 Besançon, France

^b Neurosciences intégratives et cliniques EA 481, Univ. Franche-Comté, Univ. Bourgogne Franche-Comté, F-25000 Besançon, France

^c Service de psychiatrie de l'adulte, CHRU Besançon, F-25000 Besançon, France

^d Centre Mémoire de Ressource et de Recherche de Franche-Comté, CHRU Besançon, F-25000 Besançon, France

^e Laboratoire de psychologie EA 3188, Université de Franche-Comté, Besançon, France

^f Service de neurologie, CHRU Besançon, F-25000 Besançon, France

ARTICLE INFO

Article history:

Received 6 October 2015

Revised 18 April 2016

Accepted 24 April 2016

Available online 27 April 2016

Keywords:

Musical imagery

Familiarity

Learning

Event-related potentials

EEG

ABSTRACT

The vast majority of people experience musical imagery, the sensation of reliving a song in absence of any external stimulation. Internal perception of a song can be deliberate and effortful, but also may occur involuntarily and spontaneously. Moreover, musical imagery is also involuntarily used for automatically completing missing parts of music or lyrics from a familiar song. The aim of our study was to explore the onset of musical imagery dynamics that leads to the automatic completion of missing lyrics. High-density electroencephalography was used to record the cerebral activity of twenty healthy volunteers while they were passively listening to unfamiliar songs, very familiar songs, and songs previously listened to for two weeks. Silent gaps inserted into these songs elicited a series of neural activations encompassing perceptual, attentional and cognitive mechanisms (range 100–500 ms). Familiarity and learning effects emerged as early as 100 ms and lasted 400 ms after silence occurred. Although participants reported more easily mentally imagining lyrics in familiar rather than passively learnt songs, the onset of neural mechanisms and the power spectrum underlying musical imagery were similar for both types of songs. This study offers new insights into the musical imagery dynamics evoked by gaps of silence and on the role of familiarity and learning processes in the generation of these dynamics. The automatic and effortless method presented here is a potentially useful tool to understand failure in the familiarity and learning processes of pathological populations.

© 2016 Elsevier Inc. All rights reserved.

1. Introduction

Almost everyone has already experienced the sensation of hearing a song in their mind without explicitly hearing it (Liikkanen, 2008). This internal perception can be based on a song we already know, or thinking it up as we go along. Tunes sometimes stay with us throughout the day, and are pleasant for majority of us but they can also be disturbing (Beaman & Williams, 2010). The internal recall and replay of a song engages memory mechanisms that require a substantial allocation of resources and is subsequently

cognitively demanding. The ability to recreate a musical experience when no audible sounds are present is referred to as musical imagery.

Various neuroimaging methods have been used to explore the musical imagery of individuals who deliberately imagine a piece of music or a song. Practically all studies reported similar spatial and temporal processes during the imagination of a song and its perception (Halpern & Zatorre, 1999; Linke & Cusack, 2015; Schaefer, Vlek, & Desain, 2011; Vlek, Schaefer, Gielen, Farquhar, & Desain, 2011; Zatorre, Halpern, Perry, Meyer, & Evans, 1996). The mental perception of a familiar piece of music occurs with a striking accuracy, not only the lyrics but also the tempo (Halpern, 1988; Levitin & Cook, 1996), pitch (Halpern, 1989; Janata & Paroo, 2006),

* Corresponding author at: Centre d'Investigation Clinique CIC 1431, CHRU Besançon, 2 place Saint-Jacques, F-25000 Besançon, France.

E-mail address: dgabriel@chu-besancon.fr (D. Gabriel).

timbre (Halpern, Zatorre, Bouffard, & Johnson, 2004) and emotions (Lucas, Schubert, & Halpern, 2010).

Apart from being voluntary and effortful, music imagery can also be involuntary, as commonly reported by a large proportion of the general population, who experience an internal perception of music spontaneously, while doing nothing. For example when hearing the beginning of a familiar song or of a catchy advertising jingle, it seems as if we cannot help but complete the lyrics of the song, even if we do not like it or when the advertising slogan is annoying. Pathological manifestations of this sometimes happen in schizophrenia and epilepsy, in the form of musical hallucinations. Involuntary musical imagery can be triggered by inserting gaps of noise or silence of several seconds into familiar songs (Kraemer, Macrae, Green, & Kelley, 2005; Muller et al., 2013). In this situation, and without giving any specific instructions to participants, the continuation of a song is automatically internally heard during the silence. A network of brain regions similar to the one observed during effortful musical imagery is observed, with an activation of the auditory association areas, of the dorsolateral prefrontal cortex, and the supplementary motor area (Kraemer et al., 2005).

Compared to voluntary musical imagery, one advantage to investigate spontaneous imagery with this type of design is that the onset of musical imagery is time-locked to the occurrence of silence. This offers the opportunity to explore the neural mechanisms that trigger an internal perception of music with millisecond accuracy, which was the main aim of our study. We chose to explore the temporal and spatial dynamics by means of high density EEG, whose high temporal resolution is specifically adapted to the investigation of the onset of cerebral events and whose spatial resolution allows localizing the neural generators of these events.

Our main goal was to determine when, where and how familiarity processes occurred to trigger musical imagery. We hypothesized that the unexpected silent gaps inserted into the songs would generate distinct neural events, both in the short-term and the long-term, some of these events being related to familiarity and musical imagery processes. In the short-term, we expected to detect attentional neural responses reflecting a violation of participants' prediction about the following lyrics, melody and rhythm (Vuust & Witek, 2014). Several brain potentials, such as the mismatch negativity (MMN) and P3a are known to represent preconscious, memory-based comparison processes (May & Tiitinen, 2010; Polich, 2007). Unexpected variations of pitch, timbre, or rhythm in melodies (Tervaniemi, Huotilainen, & Brattico, 2014), for example, are known to generate such potentials. Some authors advocate that musical imagery is necessary to generate these predictions about the continuation of a songs (Schaefer, 2014). This type of imagery may also be present during the perception of a song, which would explain the shared neural structures observed during music imagination and music perception. In the long term, we measured imagery mechanisms during the whole gap of silence by analyzing the EEG spectral response. During the gaps, we expected a decrease of neural synchronization in the theta (5–8 Hz) and alpha (8–12 Hz) bands for familiar songs. A similar response has previously been observed when gaps of noises were inserted into familiar songs, and was interpreted as a plausible cerebral mechanism responsible for musical hallucinations (Muller et al., 2013), a pathological manifestation of spontaneous musical imagery.

To explore further familiarity processes, the influence of different levels of familiarity on spontaneous musical imagery was also measured. In most studies using familiar content to study musical imagery, the selected pieces of music were very familiar to the listeners in order to obtain the highest number of discrepancies with unfamiliar content. Although songs running through our head are not necessarily songs heard during childhood or songs that we

know very well. They can be songs heard a few times on the radio without paying any special attention to them. It is well known that the passive exposure to music can alter the neuronal responses to musical sounds, even in nonmusicians (Koelsch, Gunter, Friederici, & Schroger, 2000; Skoe & Kraus, 2010; Wong, Perrachione, & Margulis, 2009). Consequently, we explored whether the passive exposure to music for two weeks was sufficient to induce similar spontaneous imagery mechanisms for familiar songs.

2. Methods

2.1. Participants

Twenty volunteers (7 male, 13 female, mean age = 32 ± 5 years old) participated in the experiment. All participants were right handed, nonmusicians (i.e. who did not receive any formal musical training or did so for less than 6 years), with no neurologic or psychiatric diseases, reported normal hearing and were native French speakers. All subjects gave their written consent prior to participate in the study and received compensation. None of the participants had memory deficits, as evaluated by neuropsychological evaluation. Research methods were approved by the Committee for the Protection of Human Subjects at Clinical Center Investigation of Besançon (number 14/458), where the experiment was performed.

2.2. Experimental design

The experiment consisted of three main phases: selection visit, first EEG recording, second EEG recording after a two-week learning period.

2.2.1. Selection visit

One week prior to the first data recording, participants were asked to freely choose 5 familiar songs in their native language from different singers. Five unfamiliar songs in subjects' native language were then selected by the experimenter among the original soundtracks recorded by independent or relatively unknown singers. Each of the unknown songs was carefully matched to be similar to one of the 5 familiar songs as regards musical genre and singer's gender. Once familiar and unfamiliar songs were selected, all were normalized to the same loudness level with the Replay-Gain algorithm.

2.2.2. First EEG recording

At the beginning of this session, subjects were informed about the experimental procedure that was divided into three parts: familiarity judgment of known and unknown songs, data recording and evaluation of musical imagery. After listening to a 10-s excerpt of the 10 selected songs (5 familiar and 5 unfamiliar), subjects were requested to rate each song using a visual analog scale ranging from 0 (unfamiliar) to 10 (very familiar). The knowledge of lyrics for each song could thus be checked for familiar songs (i.e. subjects were certain to know the song and thought that they might be capable of recalling most of the lyrics) as well as for unknown songs (i.e. subjects did not know any of the songs).

During the EEG recording, participants had to passively listen to the 5 familiar songs (referred to as familiar1) and the 5 unknown songs, each song being repeated twice in a random order. Songs were 2 min in length and some portions were extracted and replaced with two-second silence sections (20×2 s sections per song). These silent gaps were randomly embedded between 10 and 110 s after song onset and with a distance ranging from 2.5 to 3.5 s between sections. This resulted in 200 trials within familiar songs ($20 \text{ gaps} \times 5 \text{ familiar songs} \times 2 \text{ repetitions}$) and within

unknown songs. Silence generation and song presentation were accomplished using the E-Prime software (Psychology Software Tools Inc. Sharpsburg, PA). Audio streams embedded with gaps were played through a headphone with the volume adjusted to a comfortable listening level for each subject.

After the EEG recording, subjects were required to rate whether they were mentally completing the gaps during the experiment (0 = not at all, 10 = very well).

2.2.3. Learning period and second EEG recording

The unknown songs were then burnt onto a CD and given to each participant after the experiment. Each subject was required to listen regularly to the unknown songs for two weeks and note the frequency of song listened to per day in a self-monitoring log-book which was then collected during the second EEG recording. Then, the same experiment was repeated by using the same familiar (referred to as familiar2) and previously unfamiliar, now referred to as newly learned songs in order to assess the song learning capacity. Since the order of the songs was randomly selected, it was likely to differ from the first EEG recording. After the second EEG recording, subjects were also required to rate whether they had mentally completed the gaps during the presentation of the songs.

2.3. Data acquisition

Subjects were seated in a comfortable chair in a dark quiet testing room. EEG signals were recorded using a 256-channel Geodesic Sensor Net (Electrical Geodesics Inc.; EGI, Eugene, OR). The net is constructed to cover the skull as well as face and neck, with 20–25 mm inter-electrode distances in order to measure the arising electrical potentials from basal brain regions, thus improving/maximizing the spatial resolution of EEG. All channels were referenced to the vertex (Cz) and collected via a high impedance amplifier Net Amp 300 amplifier (Electrical Geodesics) and Net Station 4.5 software (Electrical Geodesics). Data were continuously recorded with a high-pass filter at 1 Hz, and a sampling rate at 1000 Hz. During the recording, subjects were instructed to close their eyes and remain still while passively listened to the song but they could blink and stretch as much as they wanted between each song. Subjects were not explicitly requested to complete the gaps in order to ensure that musical imagery was spontaneously/effortlessly produced. The recording time was approximately fifty minutes with a break for every 5 songs to ensure alertness and comfort of participants.

2.4. Data analysis

2.4.1. Preprocessing

Electrophysiological data were analyzed using Cartool software (version 3.55; <http://brainmapping.unige.ch/Cartool.php>). Epochs of 2500 ms were extracted from the raw data beginning 600 ms pre-silence onset and ending 1900 ms post-silence onset separately for the four conditions. A band-pass filter between 1 and 30 Hz and a notch of 50 Hz were applied to remove unwanted frequency components. The application of a baseline correction was re-considered because the real issue was to choose exactly when to apply the baseline period. We had the possibility of either applying it before silence onset (–600 to 0 ms) or toward the end of the silent period (1500–1900 ms), or using no correction at all, as advocated by some authors (see (Michel, Koenig, Brandeis, Gianotti, & Wackermann, 2009) for a discussion on baseline correction). The use of a baseline correction prior to the silence period may introduce biases due to the presentation of music and the use of a baseline correction toward the end of the silent period may conceal long-term imagery-related cognitive activities. We

decided to select the one that had been selected in a prior MEG study using a similar design (Muller et al., 2013) ranging from –600 ms to –100 ms in order to prevent any confusion over the baseline period with the silent periods. We nevertheless compared the two possible baseline corrections on a few subjects but found no real difference in the resulting signal.

Periods with visually detectable artifacts (e.g. blinks, eye movements and gross movements) were removed from the analysis. The remaining data for each subject were averaged and individual channels with artifacts were interpolated using a 3-dimensional spline algorithm (maximum: 7% interpolated electrode per subject).

2.4.2. Segmentation

To analyze event-related potentials (ERP) data, we chose not to select a priori based Time Windows (TW) of interests based on the previously reported literature. In fact, because of the unusual design of this experiment, we could not obtain any preconceived predictions regarding the time-course of ERPs obtained by the gaps. Furthermore, since conventional ERP analyses are not specifically well designed to show variations of the global electric field configuration (Michel, Seeck, & Landis, 1999), a spatiotemporal analysis that takes into account all 256 surface electrodes was carried out simultaneously. This type of procedure allows the determination of differences in electrocortical configurations between unknown, known and newly learned songs. Spatiotemporal analysis relies on a temporal segmentation based on cluster analysis (Lehmann, 1987; Michel et al., 2004), a tried and tested method for analyzing cortical networks in healthy subjects (Martinovic et al., 2014; Vocat, Pourtois, & Vuilleumier, 2008) and pathological populations (Barcelona-Lehmann et al., 2010). Time periods of stable electric configuration were first determined by segmenting together grand-mean EEG signals under the four conditions: familiar before, unfamiliar before, familiar after, newly learned songs. By testing the global field power from all electrodes, the segmentation revealed TW of interest where map topography remained stable for a certain period of time and then abruptly switches to a new configuration in which it remains stable again.

2.4.3. Regions of interest (ROIs)

Nine ROIs were selected based on the topography of ERP responses observed in previous research on auditory imagery in which they were statistically evaluated: left anterior, central anterior, right anterior, left median, central median, right median, left posterior, central posterior and right posterior (Fig. 1). Mean amplitude for each condition was computed for each TW selected on the basis of the microstate segmentation.

2.4.4. Source imaging

Intracranial source locations that accounted for different electrocortical map configurations were estimated using a distributed linear inverted solution based on a local auto-regressive average (LAURA) model (de Peralta et al., 2001). LAURA estimates three-dimensional current density distributions using a realistic head model with a solution space of 4024 nodes equally distributed within the gray matter of the Montreal Neurological Institute's average brain. Source imaging was performed during each TW of interest resulting from the segmentation. Activations obtained by source imaging were compared for each condition and each TW and thus identified diverse regional brain activities.

2.4.5. Spectral analysis

To evaluate the long-term distribution of the signal strength in the frequency domain during the entire silent period, power spectrum analysis was carried-out on each ROI using EEGLab software.

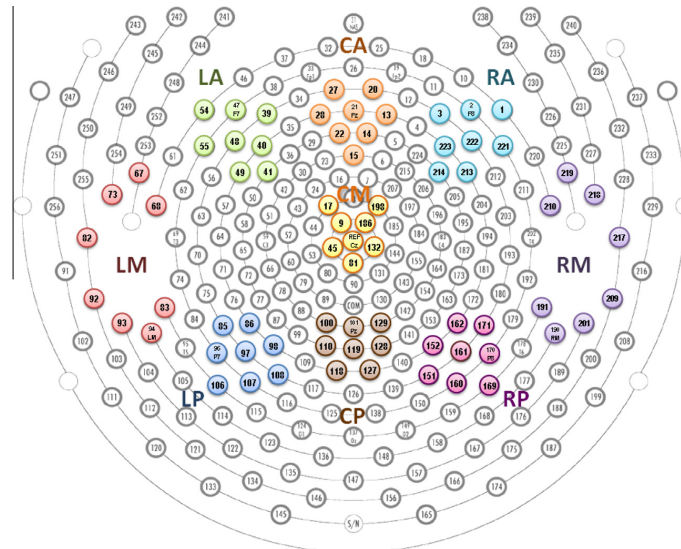


Fig. 1. Selection of the 9 regions of interest. A: Anterior, M: Median, P: Posterior; L: Left, C: Central, R: Right.

A permutation test was used for statistical evaluation. The threshold of significance used in all analyses was $p \leq 0.05$.

3. Results

3.1. Behavioral results

To verify the knowledge of lyrics for both familiar and unknown songs, familiarity judgment was first analyzed which resulted in obtaining a notion of subjects' ability to complete the gaps of silence (Table 1). Two-way repeated measures ANOVAs were performed: familiarity (2 levels: familiar and unknown songs) and learning (2 levels: before and after learning). A significant difference was found in terms of familiarity ($F(19) = 384.21$; $p < 0.0001$) between unknown and familiar songs. During the first data recording, unknown songs were not recognized by any participant compared to the familiar songs which were rated as "very familiar". By contrast, after passive listening to unknown songs twice a day for two weeks (Average = 2.39 listening per day), a significant difference was observed between unknown and learned songs during the second EEG recording, indicating a genuine learning effect in subjects. However, learned songs were still rated as less familiar than familiar songs by the participants. Similar statistical results were obtained for the subjects ability to complete the gaps ($F(19) = 123.50$; $p < 0.0001$) between unknown and familiar songs, which showed that subjects were having difficulties completing the gaps while listening to unknown songs (Table 1). They remained less capable of completing the gaps in learned songs than in familiar songs after a two-week learning phase.

Table 1
Familiarity and gap completion scores for the four categories of songs. Mean scores are presented, followed by the standard deviation.

	Familiar1	Unknown	Familiar2	Newly learned
Familiarity	8.73 ± 1.06	0.01 ± 0.05	9.01 ± 0.68	7.42 ± 1.44
Gap completion	8.53 ± 0.78	2.03 ± 1.70	8.71 ± 0.79	7.40 ± 0.94

3.2. Event-related brain potentials results

The segmentation procedure revealed 6 TW of stable topography in the 500 ms following the gap of silence (Fig. 2). The three first TWs showed clearly distinguishable ERPs under the four conditions of the study. The component on the first TW (TW1) referred to a negative component, with a latency range/an interval from 91 to 145 ms, at frontocentral sites. On the second TW (TW2), the second negative component ranged from 151 to 206 ms and was clearly evident in the central region. Then, on the third TW (TW3), a positive and centrally located component was observed, extending from 216 to 343 ms. From 350 to 510 ms a fourth TW (TW4) was also observed for all types of songs. However, for unknown songs, the segmentation revealed two additional components on the fifth TW (TW5, latency range = 372–441 ms) corresponding to a negative ERP present on central electrodes only; and on the sixth TW (TW6, latency range = 459–499 ms) which corresponded to a transition period specific to unknown songs.

To analyze these TWs, mean amplitudes of each ROI were submitted to repeated measures analysis of variance (ANOVA), repeated measures factors i.e. Familiarity (familiar/unknown), Training (before/after), Gradient (anterior/median/posterior) and Laterality (left/central/right). TW5 and TW6 was embedded into TW4, this latter window was not analyzed. A summary of the statistical results for all TWs is presented in Table 2.

Statistical analysis of TW1 only showed an interaction between Familiarity, Training, Gradient and Laterality ($F(4,76) = 3.32$, $p < 0.05$), ERP for unknown songs which were more negative on the CM ROI compared to newly learned songs ($p < 0.01$; see Fig. 3).

Analysis of TW2 revealed an interaction between Familiarity and Training ($F(1,19) = 7.78$, $p < 0.05$), unknown songs having a different pattern of activity than newly learned songs. An interaction between Familiarity, Training, and Laterality was also found ($F(2,38) = 11.7$; $p < 0.001$), revealing that there was a major difference between unknown songs and all other categories of songs on the central ROIs ($p < 0.05$ for all comparisons). Finally, an interaction between Familiarity, Training, Gradient and Laterality ($F(4,76) = 10.02$, $p < 0.001$) revealed differences between unknown

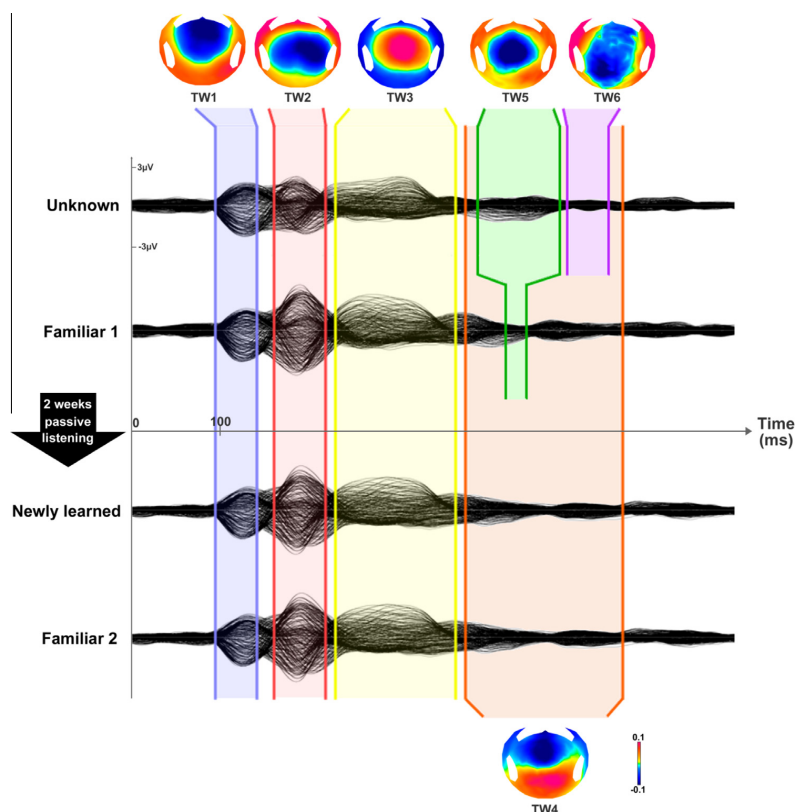


Fig. 2. Time course and topography of neural responses evoked by gaps of silence in familiar and unknown songs, as well as in familiar and newly learned (i.e., previously unknown) songs after a two-week of passive listening period. Six time windows (TW) of interest were selected after the segmentation procedure. The first three are common to all songs whereas the last three are specific to some conditions.

Table 2

Summary of the significant interactions between each type of songs in the five TWs of interest. U: Unknown songs. F1: Familiar1 songs. NL: Newly Learned songs. F2: Familiar2 songs. n.s.: not significant.

	Familiarity × training	Familiarity × training × laterality	Familiarity × training × gradient	Familiarity × training × laterality × gradient
TW1 (91–145 ms)	n.s.	n.s.	n.s.	U vs NL
TW2 (151–206 ms)	U vs NL	U vs NL U vs F1 U vs F2	n.s.	U vs NL U vs F1 U vs F2
TW3 (216–343 ms)	n.s.	n.s.	n.s.	n.s.
TW5 (372–441 ms)	n.s.	n.s.	U vs F1	U vs NL U vs F1 U vs F2
TW6 (459–499 ms)	n.s.	n.s.	U vs NL	U vs NL U vs F1 U vs F2

songs and all other categories of songs on the CM ROI ($p < 0.001$), as well as a difference between unknown songs and newly learned songs on the RM and CP ROIs ($p < 0.05$ for all), and differences between unknown songs and familiar1 songs on the LM, RM and CP ROIs ($p < 0.01$ for all).

No significant differences between conditions were found in TW3.

In TW5, an interaction between Familiarity, Training, and Gradient was found ($F(2,38) = 4.47$; $p < 0.05$), revealing differences of amplitude between unknown songs and familiar1 songs on posterior ROIs ($p < 0.05$). More specifically, an interaction between Familiarity, Training, Gradient and Laterality ($F(4,76) = 4.21$, $p < 0.01$) showed differences between unknown songs and all other types of songs on the CM and CP ROIs ($p < 0.01$ for both). Unknown

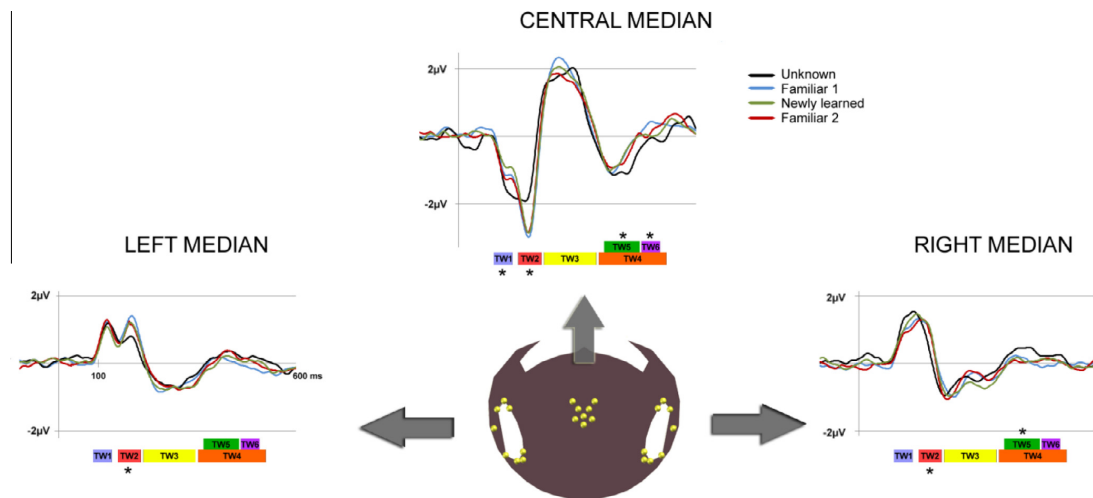


Fig. 3. Response waveforms elicited by silent gaps on the 3 median ROIs (LM, CM, RM). The 6 time windows are displayed below the ERPs of each ROI. * represent significant ERP differences ($p < 0.05$) in a specific TW between unknown songs and at least one other category of songs.

songs were also specifically different from newly learned songs on the CP ROI ($p < 0.05$) and different from familiar2 songs on RM ROI ($p < 0.01$).

Similarly to TW5, an interaction between Familiarity, Training, and Gradient ($F(2,38) = 4.47$; $p < 0.05$) was observed in TW6. Electrophysiological responses for unknown songs were different from newly learned songs on posterior ROIs ($p < 0.05$). An interaction between Familiarity, Training, Gradient and Laterality ($F(4,76) = 4.15$, $p < 0.01$) was also found and revealed differences between unknown songs and all other types of songs on the LP and CP ROIs ($p < 0.01$ for both). Differences between unknown songs and newly learned songs were also observed on the LA ($p < 0.001$), CA ($p < 0.01$), and RP ($p < 0.01$) ROIs. A specific difference between unknown and familiar1 songs was also observed on the CM ROI ($p < 0.001$).

3.3. Source localization

Source localization was performed in order to distinguish between the brain mechanisms related to each category and each TW of interest. Visually, the source localization performed in TW1 confirmed the activation of mostly temporal but also frontal regions under all conditions. Activation of frontal and temporal regions was observed in TW2 for both familiar and learned songs, with stronger activations for familiar songs compared to learned songs. Supplementary motor area (SMA) was also activated under all conditions except for unknown songs (Fig. 4). A strong activation of SMA was observed under all conditions as regards TW3. Source localization in TW4 primarily showed the activation of SMA and a slight response from prefrontal region for familiar and newly learnt songs. Although a specific cerebral activity was observed for unfamiliar songs at TW5, source localization did not show any clear differences with TW6, albeit a slightly more posterior activation of SMA. In TW5, no clear generators of cerebral activity were detected.

3.4. Spectral analysis

Power spectrum analysis was carried out over the nine ROIs in order to assess brain the mechanisms involved in long-term activ-

ities during the silent period. A significant phase difference between familiar and unfamiliar songs, predominantly between 5 and 10 Hz, corresponding to both higher theta (5–7 Hz) and lower alpha (8–10 Hz) band, was observed in the central region (Fig. 5). Note that power modulations were derived after baseline correction indicating that the differences were not simply due to a differential processing of familiar and unknown songs, but that they were specific for the period of silence. After a two-week learning phase, no significant difference was found between familiar and unknown songs in the same region (central region) in alpha, but a difference was present in theta. No difference in power spectrum was found between familiar and newly learned songs.

4. Discussion

The present study investigated the onset of cognitive processes evoked by unexpected silent gaps in familiar and passively learnt songs. Our main objective was to resolve the dynamics of activation of the cerebral structures involved during spontaneous imagery. To map functional brain areas engaged in musical imagery, we chose to use original songs instead of simplifying them into simple and artificial material such as playing notes of a piano. Although such stimuli engages different acoustic features related to both music and speech, this is a more realistic approach that reliably replicates the musical environment our brain has adapted to (Burunat, Alluri, Toiviainen, Numminen, & Brattico, 2014). One main original feature of the design presented here is that musical imagery was spontaneously evoked by the unexpected appearance of a gap of silence inserted into the song, without the previous knowledge of the participants. This paradigm is adapted for neuroimaging investigations of musical imagery since the onset of cerebral mechanisms is time-locked into the beginning of the silence. However, this type of a design conceptually differs from most of neural investigations of musical imagery where its initiation and cessation is under conscious control of participants. Behaviorally, if some association between musical expertise and voluntary musical imagery has been shown, such a relationship between voluntary and involuntary musical imagery could not be demonstrated (Weir, Williamson, & Müllensiefen, 2015). As a consequence, depending on whether musical imagery occurs in the

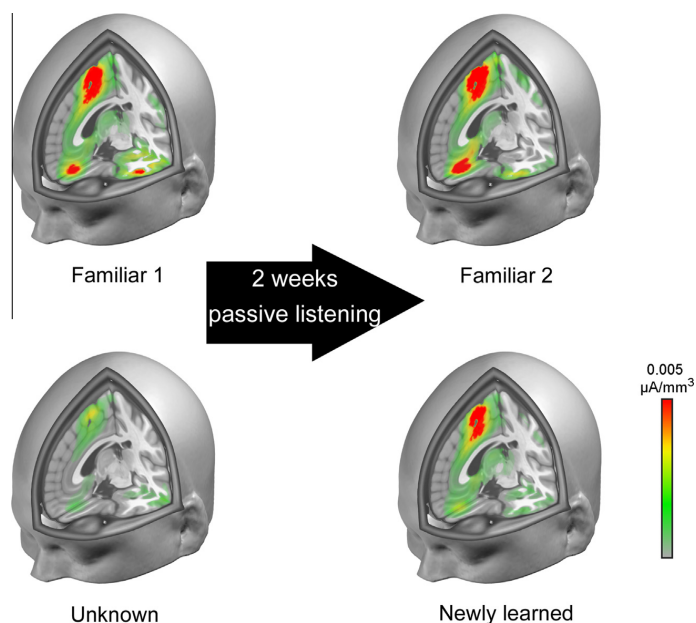


Fig. 4. Source localization of brain response in the second time windows (151–206 ms). An activation of the prefrontal cortex and the SMA is observed for familiar songs. The prefrontal cortex was less activated for newly learned songs. Unknown songs elicited only a minor activation of the SMA and prefrontal cortex.

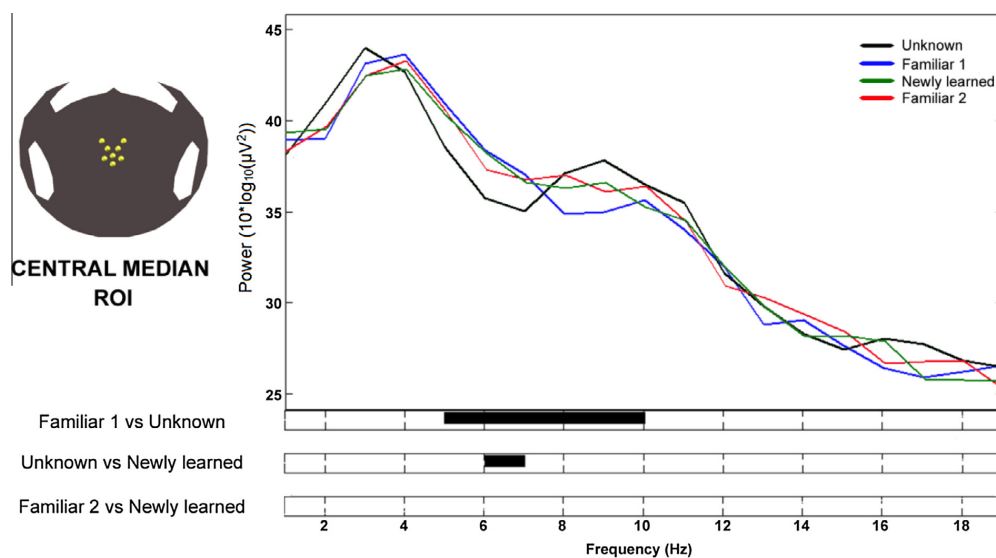


Fig. 5. Differences in the power spectrum during the period of silence in the central median ROI. Top: Representation of each curve for the four types of songs. Down: Statistical comparisons between each type of song. Significant frequency differences ($p < 0.05$) between each condition are displayed by the black bar.

presence or absence of direct volitional control, different cerebral networks may be involved.

In this study, the appearance of silent gaps triggered a whole range of cerebral activations within the first 500 ms. These activations encompass both sensory responses automatically evoked by the appearance of silence, pre-attentive and attentive responses, and more cognitive mechanisms.

4.1. Early cerebral response

In all conditions, the first brain response was evoked approximately 100 ms after the appearance of silence and had a negative distribution mostly over the frontocentral region of the scalp. Its latency and topography closely matches one of the N100 ERP, a potential involved in auditory perception, at the difference that it

was not engendered by the onset of a sound but the offset of music. The N100-Off response occurs at the cessation of sounds longer than 100 ms (Clynes, 1969) and shares many characteristics with the larger and more easily elicited onset response (Hillyard & Picton, 1978). Both -On and -Off responses appear to indicate changes in the auditory environment and thus index pre-attentive auditory processing underlying stimulus detection. Whether the neural populations responsible for the N100 responses occurring at the onset or the offset of a sound overlap or not remains debated (Noda et al., 1998; Wakai, Lutter, Chen, & Maier, 2007). However, we show here that the N100-Off response is dependent on familiarity processes, contrary to what has been previously observed with the N100-On (Purhonen, Paakkonen, Ypparila, Lehtonen, & Karhu, 2001). Since this effect was only noticeable between unknown and newly learned songs on the CM ROI, this will have to be further investigated in future studies.

4.2. Attentional mechanisms

4.2.1. Mismatch negativity (MMN)

The negative central distribution occurring around 150–200 ms after the appearance of silence is typical of a MMN. This assumption is supported by the presence of prefrontal and auditory generators that are typical for the elicitation of this potential (Giard, Perrin, Pernier, & Bouchet, 1990; Gomot, Giard, Roux, Barthelemy, & Bruneau, 2000; Maess, Jacobsen, Schroger, & Friederici, 2007). While MMN is usually reported to signal a mismatch between sensory input and stored memory representations (Folstein & Van Petten, 2008), the elicitation of this ERP can also be based on a purely imagined memory trace (Herholz, Lappe, Knief, & Pantev, 2008, 2009). Here, the MMN was obtained by silences interspersed within the songs, and was elicited by the unexpected absence either of lyrics, melody, rhythm, or all these parameters together. These results are consistent with theories of predictive coding, which suppose that during the song the brain was continuously extracting structural regularities to make predictions of what would be heard next (Friston, 2002; Vuust, Ostergaard, Pallesen, Bailey, & Roepstorff, 2009; Vuust & Witek, 2014). The violation of these predictions by the appearance of silence generated an error response from the brain.

Interestingly, we found that familiarity affected the MMN to some extent since its amplitude was significantly smaller for unknown songs compared to other songs. This differs from previous results obtained with melody alone (Neuloh & Curio, 2004), although it does support findings on speech processing where familiar phonemes (Beauchemin et al., 2006; Naatanen, 2001) and prosody (Ylinen, Strelnikov, Huotilainen, & Naatanen, 2009) typically leads to enhanced MMN amplitude. A similar effect was also reported in the neural investigations of rhythmic incongruence, where an enhanced MMN was observed in expert musicians compared to nonmusicians (Vuust et al., 2009). Compared to unfamiliar songs, listening to familiar songs may elicit stronger expectancies related to its rhythmic progression.

Source localization showed a stronger activation in the prefrontal area and the SMA for familiar compared to unknown songs, two regions involved in the imagery of familiar songs (Herholz, Halpern, & Zatorre, 2012). These regions are activated during the perception (Pereira et al., 2011) but even more during illusion (Halpern & Zatorre, 1999; Kraemer et al., 2005) of familiar songs, underlying the increased processing demands of imagery (Herholz et al., 2012). The SMA region is also considered as a potential generator of the MMN (Huang et al., 2010).

4.2.2. P3a

The positive frontocentral distribution observed around 300 ms under all conditions is similar to a P3a, an ERP evoked in response

to infrequent or novel stimuli (Squires, Squires, & Hillyard, 1975). This is supported by source localization, whose generators were mostly localized around the SMA, corroborating previous observations (Rektor et al., 2003; Rusiniak et al., 2013). In the musical context, a P3a was reported when a sequence of chords is followed by harmonically improbable chords (Janata, 1995). The response is independent from familiarity and learning, an absence of effect previously observed with melody (Neuloh & Curio, 2004). However, this differs from speech studies that found differences in P3a between unfamiliar and familiar voices (Beauchemin et al., 2006), as well as between unfamiliar and the subject's own voice (Graux et al., 2013), and in rhythmic incongruity between musicians and nonmusicians (Vuust et al., 2009).

It is puzzling that whereas the MMN and the P3a are both related to attentional processes and share similar generators, it is the earlier and not the later ERP that is related to familiarity, and that the P3a process occurs independently from the amplitude of the MMN response. Although the P3a is in many ways related to the MMN, it can be seen as a correlate of the processes that accompanies the redirection of voluntary attention, as detected by the MMN (Chennu & Bekinschtein, 2012). The reorientation of attention thus occurs whether the song is familiar or not.

4.3. Cognitive responses

Between 350 ms and 500 ms, the segmentation procedure revealed a long period of stability for all familiar and newly learned songs but not for unknown songs. This TW is reflected by negative frontocentral and positive parietal brain activity. For unknown songs, the beginning of this TW showed a similar activity but was quickly replaced by two distinct TWs, one with a central negativity and the other that is more a transition with a less defined cortical activity.

Processes occurring between 350 and 500 ms are known to be associated with memory performance. A number of neurocognitive studies have been reported on correlations between a component of this TW, the P300, and the memory processes and learning (Friedman & Johnson, 2000; Jongsma, Gerrits, van Rijn, Quiroga, & Maes, 2012). Several studies have found a relationship between recall probability and increased P300 amplitude (Curran & Cleary, 2003; Karis, Druckman, Lissak, & Donchin, 1984; Paller, McCarthy, & Wood, 1988; Rushby, Barry, & Johnstone, 2002). A larger P300 is observed when stimuli are well encoded and stored in memory compared to less successfully encoded and stored stimuli (Polich, 2007).

Also, the frontocentral negative topography observed in this TW could well be related to a N400 familiarity effect that has previously been observed elsewhere (Daltrozzo, Tillmann, Platel, & Schon, 2010). However, no difference was observed between familiar and newly learned songs. One may hypothesize that either newly learned songs were familiar enough to engender the maximum activation in this TW compared to moderately familiar music used in previous study (Daltrozzo et al., 2010), or that the association of words has potentialized the strength of the response, even for less familiar songs.

It is noteworthy that similar to some other TW, prefrontal cortex and SMA were activated in this TW. The activation of these structures confirms previous results obtained in spontaneous musical imagery with fMRI in a similar paradigm (Kraemer et al., 2005). For example, the SMA was activated in all TWs except the first and the last one for unfamiliar songs. This structure is known to be involved in various auditory processes such as attention and working memory (Rama et al., 2001; Wager & Smith, 2003; Zhang, Leung, & Johnson, 2003), in mental imagery (Halpern, 2001; Halpern & Zatorre, 1999; Halpern et al., 2004; Kraemer et al., 2005), as well as in the perception (Pereira et al., 2011) and antic-

ipation (Leaver, Van Lare, Zielinski, Halpern, & Rauschecker, 2009) of familiar music. Since all these functions contribute to the internal perception of music, the activated SMA in each TW may have played different roles in the generation of involuntary musical imagery and will need to be further investigated.

Compared to other songs, unknown songs elicited a clear central negative ERP at approximately 400 ms that might reflect a reaction to the unexpected and inappropriate appearance of silence. This would define the typical N400, an electrophysiological index of the processing of meaningful information. In this case it would be related to the increased difficulty to complete the lyrics, even if it has an access to some information. However, the fact that the same topography as TW5 was present for familiar songs for a short duration (24 ms) suggests that other processes may be involved, such as the familiarization to the task, but this has to be further explored. For unknown songs, the topography of the sixth TW typically reflects a transition phase related to impossibility to further complete the gap of silence.

4.4. Whole-silence cerebral activity

The transient activity generated by the appearance of silence lasted only 500 ms and did not provide any information about the long-term musical imagery dynamics engaged during this period. In situations where a continuous auditory stimulus is employed, such as listening to a song, the EEG is primarily analyzed in the frequency domain (Sammler, Grigutsch, Fritz, & Koelsch, 2007). The same applies here to understand the mechanisms engaged during the whole period of silence. Our assumption was that mental song completion would be associated with a sustained and stable pattern of neural activity. The central ROI exhibited alpha power desynchronization observed as well as theta power synchronization when gaps of silence were embedded in familiar songs as compared to unknown songs. This confirms previous results obtained during gaps of noise with MEG in healthy participants and intracranially recorded activity in epilepsy patients (Muller et al., 2013). In these patients, the alpha power reduction was found to originate from both auditory regions, essentially from the secondary auditory cortex, and non-auditory regions such as the inferior and medial frontal cortex. Reduced alpha power in the auditory cortex is similar to modulations occurring during auditory perception (Weisz, Hartmann, Muller, Lorenz, & Obleser, 2011), confirming that the processes involved in spontaneous imagery and perception share similar mechanisms. A desynchronization in alpha power is also present during cognitive processing invoked by task conditions requiring attention and memory (Boiten, Sergeant, & Geuze, 1992; Pfurtscheller & Klimesch, 1991, 1992; Sergeant, Geuze, & van Winsum, 1987). A recent study showed that the frequency of spontaneous imagery was related to cortical thickness in several regions and notably temporal regions, suggesting the important role of these structures in the generation of this phenomenon (Farrugia, Jakubowski, Cusack, & Stewart, 2015). Although it has been suggested that a reduction in an alpha band is a basis for the generation of auditory percepts (Muller et al., 2013), we did not observe any differences in alpha modulations between unfamiliar and newly learned songs. This suggests that musical imagery for newly learned songs was not as vivid as for familiar songs, as confirmed by the use of various questionnaires.

The increase in theta power found between familiar and unfamiliar songs confirms a previous observation made with invasive electrophysiological recordings (Muller et al., 2013). According to that study, the increase was related to the activation of memory-related regions such as the medial temporal lobe, a structure that communicates with other regions through phase synchronization in low frequency bands (Fuentemilla, Penny, Cashdollar, Bunzeck,

& Duzel, 2010; Lega, Jacobs, & Kahana, 2012). It has thus been suggested that theta modulations permit communication between the auditory cortex and medial temporal structures, allowing the gathering of relevant song information to effectively perform imagery (Muller et al., 2013).

5. Conclusion

In summary, the present study aimed to elucidate the neural dynamics of spontaneous musical imagery, more specifically when, where and how familiarity triggers the mental representation of songs. The design used, which automatically induces an internal musical experience by inserting precisely time-locked gaps of silence, allowed exploring the spatiotemporal dynamics underlying the mental representation of songs. Among the series of neural activations occurring in range 100–500 ms and encompassing perceptual, attentional and cognitive mechanisms, we showed that familiarity mechanisms emerged under two different forms that may reflect two distinct types of imagery (Schaefer, 2014). The first response corresponds to event-related potentials of varying amplitude occurring as early as 100 ms after the beginning of silence. In the context of predictive coding, we consider the reduction of amplitude of the MMN as a reduced capacity to make accurate predictions about the continuation song. For unknown songs, one can hypothesize that only harmonic and rhythmic predictions were possible but future studies taking into account the multi-leveled structure of music are necessary to verify such a hypothesis. Whether predictive coding is part of an imagery process or not is still discussed (Schaefer, 2014), but other neuroimaging studies have shown that the neural structures involved in the generation of the prediction process (SMA, prefrontal and temporal auditory regions) play a considerable role in musical imagery. However, it is important to point out that the paradigm used in the present study is conceptually different to the majority of studies on musical imagery where participants are explicitly requested to conjure up a musical sequence in their mind. In that respect, special care has to be taken when wishing to extend the neural dynamics we observed to voluntary imagery investigations.

The second neural response to familiarity was an increase in theta power during the whole silence period. The modulations of this frequency band are known to reflect a communication between the auditory cortex and medial temporal structures, which is essential to mentally retrieve missing lyrics in familiar songs. Interestingly, we showed that the neural familiarity mechanisms were similar between familiar songs and songs passively learned for two weeks. This shows the efficacy of passive learning processes in encoding musical information and may partly explain why a catchy piece of music can continually repeat through a person's mind after being heard a few times on radio before (Liikkanen, 2008).

If the musical design presented here was devoted to explore the neural dynamics of familiarity in healthy subjects, we firmly believe that it could also benefit pathological populations. Indeed, this method of investigating the neural responses to familiarity and learning is effortless, automatic, and does not require giving any instructions, which is often a critical aspect in the evaluation of mnemonic mechanisms in pathological populations. Therefore, this could be a useful tool applied to evaluate the integrity of musical memory in unresponsive populations with suspected memory disorders

Acknowledgment

This work was supported by a grant from the University Hospital of Besançon (Internal Project named APICHU in 2014: reference

no. API/2014/51). The authors are grateful to Richard Medeiros-Medical Editor of Medical Editing International for editing the manuscript.

Appendix A. Supplementary material

Supplementary data associated with this article can be found, in the online version, at <http://dx.doi.org/10.1016/j.nlm.2016.04.011>.

References

- Barcellona-Lehmann, S., Morand, S., Bindschaedler, C., Nahum, L., Gabriel, D., & Schneider, A. (2010). Abnormal cortical network activation in human amnesia: A high-resolution evoked potential study. *Brain Topography*, 23(1), 72–81.
- Beaman, C. P., & Williams, T. I. (2010). Earworms (stuck song syndrome): Towards a natural history of intrusive thoughts. *British Journal of Psychology*, 101(Pt 4), 637–653.
- Beauchemin, M., De Beaumont, L., Vannasing, P., Turcotte, A., Arcand, C., Belin, P., et al. (2006). Electrophysiological markers of voice familiarity. *European Journal of Neuroscience*, 23(11), 3081–3086.
- Boiten, F., Sergeant, J., & Geuze, R. (1992). Event-related desynchronization: The effects of energetic and computational demands. *Electroencephalography and Clinical Neurophysiology*, 82(4), 302–309.
- Burunat, I., Alluri, V., Toiviainen, P., Numminen, J., & Brattico, E. (2014). Dynamics of brain activity underlying working memory for music in a naturalistic condition. *Cortex*, 57, 254–269.
- Chennu, S., & Bekinschtein, T. A. (2012). Arousal modulates auditory attention and awareness: Insights from sleep, sedation, and disorders of consciousness. *Frontiers in Psychology*, 3, 65.
- Clynes, M. (1969). *The dynamics of R-M function*. Washington, DC: U.S. Government Printing Office.
- Curran, T., & Cleary, A. M. (2003). Using ERPs to dissociate recollection from familiarity in picture recognition. *Brain Research. Cognitive Brain Research*, 15(2), 191–205.
- Daltrozzo, J., Tillmann, B., Platel, H., & Schon, D. (2010). Temporal aspects of the feeling of familiarity for music and the emergence of conceptual processing. *Journal of Cognitive Neuroscience*, 22(8), 1754–1769.
- de Peralta, G., Menendez, R., Gonzalez Andino, S., Lantz, G., Michel, C. M., & Landis, T. (2001). Noninvasive localization of electromagnetic epileptic activity. I. Method descriptions and simulations. *Brain Topography*, 14(2), 131–137.
- Farrugia, N., Jakubowski, K., Cusack, R., & Stewart, L. (2015). Tunes stuck in your brain: The frequency and affective evaluation of involuntary musical imagery correlate with cortical structure. *Consciousness and Cognition*, 35, 66–77.
- Folstein, J. R., & Van Petten, C. (2008). Influence of cognitive control and mismatch on the N2 component of the ERP: A review. *Psychophysiology*, 45(1), 152–170.
- Friedman, D., & Johnson, R. Jr. (2000). Event-related potential (ERP) studies of memory encoding and retrieval: A selective review. *Microscopy Research and Technique*, 51(1), 6–28.
- Friston, K. (2002). Beyond phrenology: What can neuroimaging tell us about distributed circuitry? (0147–006X (Print)).
- Fuentemilla, L., Penny, W. D., Cashdollar, N., Bunzeck, N., & Düzel, E. (2010). Theta-coupled periodic replay in working memory. *Current Biology*, 20(7), 606–612.
- Giard, M. H., Perrin, F., Pernier, J., & Bouchet, P. (1990). Brain generators implicated in the processing of auditory stimulus deviance: A topographic event-related potential study. *Psychophysiology*, 27(6), 627–640.
- Gomot, M., Giard, M. H., Roux, S., Barthelemy, C., & Bruneau, N. (2000). Maturation of frontal and temporal components of mismatch negativity (MMN) in children. *NeuroReport*, 11(14), 3109–3112.
- Graux, J., Gomot, M., Roux, S., Bonnet-Brilhault, F., Camus, V., & Bruneau, N. (2013). My voice or yours? An electrophysiological study. *Brain Topography*, 26(1), 72–82.
- Halpern, A. R. (1988). Mental scanning in auditory imagery for songs. *Journal of Experimental Psychology. Learning, Memory, and Cognition*, 14(3), 434–443.
- Halpern, A. R. (1989). Memory for the absolute pitch of familiar songs. *Memory & Cognition*, 17(5), 572–581.
- Halpern, A. R. (2001). Cerebral substrates of musical imagery. *Annals of the New York Academy of Sciences*, 930, 179–192.
- Halpern, A. R., & Zatorre, R. J. (1999). When that tune runs through your head: A PET investigation of auditory imagery for familiar melodies. *Cerebral Cortex*, 9(7), 697–704.
- Halpern, A. R., Zatorre, R. J., Bouffard, M., & Johnson, J. A. (2004). Behavioral and neural correlates of perceived and imagined musical timbre. *Neuropsychologia*, 42(9), 1281–1292.
- Herholz, S. C., Halpern, A. R., & Zatorre, R. J. (2012). Neuronal correlates of perception, imagery, and memory for familiar tunes. *Journal of Cognitive Neuroscience*, 24(6), 1382–1397 [1530–8898 (Electronic)].
- Herholz, S. C., Lappe, C., Knief, A., & Pantev, C. (2008). Neural basis of music imagery and the effect of musical expertise. *European Journal of Neuroscience*, 28(11), 2352–2360.
- Herholz, S. C., Lappe, C., Knief, A., & Pantev, C. (2009). Imagery mismatch negativity in musicians. *Annals of the New York Academy of Sciences*, 1169, 173–177.
- Hillyard, S. A., & Picton, T. W. (1978). On and off components in the auditory evoked potential. *Perception & Psychophysics*, 24(5), 391–398.
- Huang, M. X., Lee, R. R., Gaa, K. M., Song, T., Harrington, D. L., Loh, C., et al. (2010). Somatosensory system deficits in schizophrenia revealed by MEG during a median-nerve oddball task. *Brain Topography*, 23(1), 82–104.
- Janata, P. (1995). ERP measures assay the degree of expectancy violation of harmonic contexts in music. *Journal of Cognitive Neuroscience*, 7(2), 153–164.
- Janata, P., & Paroo, K. (2006). Acuity of auditory images in pitch and time. *Perception & Psychophysics*, 68(5), 829–844.
- Jongasma, M. L., Gerrits, N. J., van Rijn, C. M., Quiroga, R. Q., & Maes, J. H. (2012). Event related potentials to digit learning: Tracking neurophysiologic changes accompanying recall performance. *International Journal of Psychophysiology*, 85(1), 41–48.
- Karis, D., Druckman, D., Lissak, R., & Donchin, E. (1984). A psychophysiological analysis of bargaining. ERPs and facial expressions. *Annals of the New York Academy of Sciences*, 425, 230–235.
- Koelsch, S., Gunter, T., Friederici, A. D., & Schroger, E. (2000). Brain indices of music processing: “nonmusicians” are musical. *Journal of Cognitive Neuroscience*, 12(3), 520–541.
- Kraemer, D. J., Macrae, C. N., Green, A. E., & Kelley, W. M. (2005). Musical imagery: Sound of silence activates auditory cortex. *Nature*, 434(7030), 158.
- Leaver, A. M., Van Lare, J., Zielinski, B., Halpern, A. R., & Rauschecker, J. P. (2009). Brain activation during anticipation of sound sequences. *Journal of Neuroscience*, 29(8), 2477–2485.
- Lega, B. C., Jacobs, J., & Kahana, M. (2012). Human hippocampal theta oscillations and the formation of episodic memories. *Hippocampus*, 22(4), 748–761.
- Lehmann, D. (1987). Principles of spatial analysis. *Methods of analysis of brain electrical and magnetic signals* (Vol. 1). Amsterdam: Elsevier.
- Levitin, D. J., & Cook, P. R. (1996). Memory for musical tempo: Additional evidence that auditory memory is absolute. *Perception & Psychophysics*, 58(6), 927–935.
- Liikkanen, L. A. (2008). In *Music in everymind: Commonality of involuntary musical imagery paper presented at the proceedings of the 10th international conference on music perception and cognition (ICMPC10)*. Sapporo, Japan.
- Linke, A. C., & Cusack, R. (2015). Flexible information coding in human auditory cortex during perception, imagery, and STM of complex sounds. *Journal of Cognitive Neuroscience*, 27(7), 1322–1333.
- Lucas, B. J., Schubert, E., & Halpern, A. R. (2010). Perception of emotion in sounded and imagined music. *Music Perception: An Interdisciplinary Journal*, 27(5), 399–412.
- Maess, B., Jacobsen, T., Schroger, E., & Friederici, A. D. (2007). Localizing pre-attentive auditory memory-based comparison: Magnetic mismatch negativity to pitch change. *NeuroImage*, 37(2), 561–571.
- Martinovic, J., Jones, A., Christiansen, P., Rose, A. K., Hogarth, L., & Field, M. (2014). Electrophysiological responses to alcohol cues are not associated with Pavlovian-to-instrumental transfer in social drinkers. *PLoS One*, 9(4), e94605.
- May, P. J., & Tiitinen, H. (2010). Mismatch negativity (MMN), the deviance-elicited auditory deflection, explained. *Psychophysiology*, 47(1), 66–122.
- Michel, C. M., Koenig, T., Brandeis, D., Gianotti, L. R. R., & Wackermann, J. (2009). *Electrical neuroimaging*. Cambridge: Cambridge University Press.
- Michel, C. M., Murray, M. M., Lantz, G., Gonzalez, S., Spinelli, L., & Grave de Peralta, R. (2004). EEG source imaging. *Clinical Neurophysiology*, 115(10), 2195–2222.
- Michel, C. M., Seec, M., & Landis, T. (1999). Spatiotemporal dynamics of human cognition. *News in Physiological Sciences*, 14, 206–214.
- Muller, N., Keil, J., Obleser, J., Schulz, H., Grunwald, T., Bernays, R. L., et al. (2013). You can't stop the music: Reduced auditory alpha power and coupling between auditory and memory regions facilitate the illusory perception of music during noise. *NeuroImage*, 79, 383–393.
- Naatanen, R. (2001). The perception of speech sounds by the human brain as reflected by the mismatch negativity (MMN) and its magnetic equivalent (MMNm). *Psychophysiology*, 38(1), 1–21.
- Neuloh, G., & Curio, G. (2004). Does familiarity facilitate the cortical processing of music sounds? *NeuroReport*, 15(16), 2471–2475.
- Noda, K., Tonoike, M., Doi, K., Koizuka, I., Yamaguchi, M., Seo, R., et al. (1998). Auditory evoked off-response: Its source distribution is different from that of on-response. *NeuroReport*, 9(11), 2621–2625.
- Paller, K. A., McCarthy, G., & Wood, C. C. (1988). ERPs predictive of subsequent recall and recognition performance. *Biological Psychology*, 26(1–3), 269–276.
- Pereira, C. S., Teixeira, J., Figueiredo, P., Xavier, J., Castro, S. L., & Brattico, E. (2011). Music and emotions in the brain: Familiarity matters. *PLoS One*, 6(11), e27241.
- Pfurtscheller, G., & Klimesch, W. (1991). Event-related desynchronization during motor behavior and visual information processing. *Electroencephalography and Clinical Neurophysiology. Supplement*, 42, 58–65.
- Pfurtscheller, G., & Klimesch, W. (1992). Functional topography during a visuo-verbal judgment task studied with event-related desynchronization mapping. *Journal of Clinical Neurophysiology*, 9(1), 120–131.
- Polich, J. (2007). Updating P300: An integrative theory of P3a and P3b. *Clinical Neurophysiology*, 118(10), 2128–2148.
- Purhonen, M., Paakkonen, A., Ypparila, H., Lehtonen, J., & Karhu, J. (2001). Dynamic behavior of the auditory N100 elicited by a baby's cry. *International Journal of Psychophysiology*, 41(3), 271–278.
- Rama, P., Martinkauppi, S., Linnankoski, I., Koivisto, J., Aronen, H. J., & Carlson, S. (2001). Working memory of identification of emotional vocal expressions: An fMRI study. *NeuroImage*, 13(6 Pt 1), 1090–1101.
- Rektor, I., Kaijovsky, P., Bares, M., Brazdil, M., Streitova, H., Klajblava, H., et al. (2003). A SEEG study of ERP in motor and premotor cortices and in the basal ganglia. *Clinical Neurophysiology*, 114(3), 463–471.

- Rushby, J. A., Barry, R. J., & Johnstone, S. S. (2002). Event-related potential correlates of serial-position effects during an elaborative memory test. *International Journal of Psychophysiology*, *46*(1), 13–27.
- Rusiniak, M., Lewandowska, M., Wolak, T., Pluta, A., Milner, R., Ganc, M., et al. (2013). A modified oddball paradigm for investigation of neural correlates of attention: A simultaneous ERP-fMRI study. *Magnetic Resonance Materials in Physics, Biology and Medicine*, *26*(6), 511–526.
- Sammler, D., Grigutsch, M., Fritz, T., & Koelsch, S. (2007). Music and emotion: Electrophysiological correlates of the processing of pleasant and unpleasant music. *Psychophysiology*, *44*(2), 293–304.
- Schaefer, R. S. (2014). Mental representations in musical processing and their role in action-perception loops. *Empirical Musicology Review*, *9*(3–4), 161–176.
- Schaefer, R. S., Vlek, R. J., & Desain, P. (2011). Music perception and imagery in EEG: Alpha band effects of task and stimulus. *International Journal of Psychophysiology*, *82*(3), 254–259.
- Sergeant, J., Geuze, R., & van Winsum, W. (1987). Event-related desynchronization and P300. *Psychophysiology*, *24*(3), 272–277.
- Skoe, E., & Kraus, N. (2010). Hearing it again and again: On-line subcortical plasticity in humans. *PLoS One*, *5*(10), e13645.
- Squires, N. K., Squires, K. C., & Hillyard, S. A. (1975). Two varieties of long-latency positive waves evoked by unpredictable auditory stimuli in man. *Electroencephalography and Clinical Neurophysiology*, *38*(4), 387–401.
- Tervaniemi, M., Huotilainen, M., & Brattico, E. (2014). Melodic multi-feature paradigm reveals auditory profiles in music-sound encoding. *Frontiers in Human Neuroscience*, *8*, 496.
- Vlek, R. J., Schaefer, R. S., Gielen, C. C., Farquhar, J. D., & Desain, P. (2011). Shared mechanisms in perception and imagery of auditory accents. *Clinical Neurophysiology*, *122*(8), 1526–1532.
- Vocat, R., Pourtois, G., & Vuilleumier, P. (2008). Unavoidable errors: A spatio-temporal analysis of time-course and neural sources of evoked potentials associated with error processing in a speeded task. *Neuropsychologia*, *46*(10), 2545–2555.
- Vuust, P., Ostergaard, L., Pallesen, K. J., Bailey, C., & Roepstorff, A. (2009). Predictive coding of music-brain responses to rhythmic incongruity. *Cortex*, *45*, 80–92 [0010-9452 (Print)].
- Vuust, P., & Witek, M. A. (2014). Rhythmic complexity and predictive coding: A novel approach to modeling rhythm and meter perception in music. *Frontiers in Psychology*, *5* [1664-1078 (Electronic)].
- Wager, T. D., & Smith, E. E. (2003). Neuroimaging studies of working memory: A meta-analysis. *Cognitive, Affective, & Behavioral Neuroscience*, *3*(4), 255–274.
- Wakai, R. T., Lutter, W. J., Chen, M., & Maier, M. M. (2007). On and Off magnetic auditory evoked responses in early infancy: A possible marker of brain immaturity. *Clinical Neurophysiology*, *118*(7), 1480–1487.
- Weir, G., Williamson, V. J., & Müllensiefen, D. (2015). Increased involuntary musical mental activity is not associated with more accurate voluntary musical imagery. *Psychomusicology: Music, Mind, and Brain*, *25*(1), 48–57. <http://dx.doi.org/10.1037/pmu0000076>.
- Weisz, N., Hartmann, T., Müller, N., Lorenz, I., & Obleser, J. (2011). Alpha rhythms in audition: Cognitive and clinical perspectives. *Frontiers in Psychology*, *2*, 73.
- Wong, P. C., Perrachione, T. K., & Margulis, E. H. (2009). Effects of asymmetric cultural experiences on the auditory pathway: Evidence from music. *Annals of the New York Academy of Sciences*, *1169*, 157–163.
- Ylinen, S., Strelnikov, K., Huotilainen, M., & Naatanen, R. (2009). Effects of prosodic familiarity on the automatic processing of words in the human brain. *International Journal of Psychophysiology*, *73*(3), 362–368.
- Zatorre, R. J., Halpern, A. R., Perry, D. W., Meyer, E., & Evans, A. C. (1996). Hearing in the mind's ear: A PET investigation of musical imagery and perception. *Journal of Cognitive Neuroscience*, *8*(1), 29–46.
- Zhang, J. X., Leung, H. C., & Johnson, M. K. (2003). Frontal activations associated with accessing and evaluating information in working memory: An fMRI study. *Neuroimage*, *20*(3), 1531–1539.

ANNEXE B : ARTICLE CORRESPONDANT AU CHAPITRE 4

ARTICLE IN PRESS

International Journal of Psychophysiology xxx (xxxx) xxx–xxx



ELSEVIER

Contents lists available at ScienceDirect

International Journal of Psychophysiology

journal homepage: www.elsevier.com/locate/ijpsycho

Evidence for a neural signature of musical preference during silence

Coralie Joucla^{a,b}, Magali Nicolier^{a,b,c}, Julie Giustiniani^{a,b,c}, Gaëlle Brunotte^a, Nicolas Noiret^{d,e},
 Julie Monnin^{a,b,c}, Eloi Magnin^{d,f}, Lionel Pazart^{a,b}, Thierry Moulin^{a,b,f}, Emmanuel Haffen^{a,b,c},
 Pierre Vandel^{a,b,c,d}, Damien Gabriel^{a,b,*}

^a Centre d'Investigation Clinique-Innovation Technologique CIC-IT 1431, Inserm, CHRU Besançon, F-25000 Besançon, France

^b Neurosciences intégratives et cliniques EA 481, Univ. Franche-Comté, Univ. Bourgogne Franche-Comté, F-25000 Besançon, France

^c Service de psychiatrie de l'adulte, CHRU Besançon, F-25000 Besançon, France

^d Centre Mémoire de Ressource et de Recherche de Franche-Comté, CHRU Besançon, F-25000 Besançon, France

^e Laboratoire de psychologie EA 3188, Université de Franche-Comté, F-25000 Besançon, France

^f Service de neurologie, CHRU Besançon, F-25000 Besançon, France

ARTICLE INFO

Keywords:

Musical preference
 Silence
 Auditory imagery
 High density EEG

ABSTRACT

One of the most basic and person-specific affective responses to music is liking. The present investigation sought to determine whether liking was preserved during spontaneous auditory imagery. To this purpose, we inserted two-second silent intervals into liked and disliked songs, a method known to automatically recreate a mental image of these songs. Neural correlates of musical preference were measured by high-density electroencephalography in twenty subjects who had to listen to a set of five pre-selected unknown songs the same number of times for two weeks. Time frequency analysis of the two most liked and the two most disliked songs confirmed the presence of neural responses related to liking. At the beginning of silent intervals (400–900 ms and 1000–1300 ms), significant differences in theta activity were originating from the inferior frontal and superior temporal gyrus. These two brain structures are known to work together to process various aspects of music and are also activated when measuring liking while listening to music. At the end of silent intervals (1400–1900 ms), significant alpha activity differences originating from the insula were observed, whose exact role remains to be explored. Although exposure was controlled for liked and disliked songs, liked songs were rated as more familiar, underlying the strong relationship that exists between liking, exposure, and familiarity.

1. Introduction

Sometimes it's the space between the notes that makes all the difference in music. Silence in music conveys information of major importance that was efficiently used by composers like Mozart who mentioned that “music is not in the notes, but in the silence between”. Some authors have suggested that high-level musical information, such as familiarity, is also present during silence: “When you listen to Mozart, the silence that follows is still Mozart” (Guitry, 1947).

In the early 2000s, neuroimaging studies confirmed that familiarity was processed during silences, even when these silences were artificially inserted into music (Kraemer et al., 2005; Gabriel et al., 2016). When inserted in familiar music, these silences automatically convey the sensation of hearing a song internally without any audible sound, in continuation of what had just been heard before. This is referred to as spontaneous auditory imagery. As a consequence of spontaneous imagery, a neural activity within regions of secondary auditory cortex, of

the dorsolateral prefrontal cortex and of the supplementary motor area is generated (Kraemer et al., 2005). In electroencephalography (EEG), the response to familiar music is reflected by an increase in theta power during the whole silence period (Gabriel et al., 2016). Interestingly, the neural familiarity mechanisms occurring during auditory imagery are similar between familiar songs and songs passively learned for two weeks, showing the efficiency of passive learning processes in encoding musical information (Gabriel et al., 2016).

The present investigation sought to determine whether aesthetic musical judgement is also present during silence in music. When we like listening to Mozart, do we still like Mozart during the silences in its music? Music is a complex stimulus, and its affective dimension is an integral part of the artistic and aesthetic experience, and the hedonic appreciation of music, referred to as music liking, involves activation of areas responsible for the feeling of pleasure (Blood and Zatorre, 2001). The hypothesis of a preservation of liking during auditory imagery is based on the important influence of familiarity on liking: becoming

* Corresponding author at: Clinical Investigation Center, Bâtiment Saint-Joseph, CHRU Saint Jacques, 2 place Saint Jacques, F-25 030 Besançon, France.
 E-mail address: dgabriel@chu-besancon.fr (D. Gabriel).

<https://doi.org/10.1016/j.ijpsycho.2018.02.007>

Received 23 August 2017; Received in revised form 16 February 2018; Accepted 18 February 2018
 0167-8760/ © 2018 Elsevier B.V. All rights reserved.

more familiar with a particular piece of music increases the subject's liking ratings for it (Pereira et al., 2011). Familiarity and liking being two high level information processes often co-occurring when listening to music (Stalinski and Schellenberg, 2013), they may as well co-occur during musical imagery. A second argument supporting that liking may be maintained is based on the fact that emotions are preserved during auditory imagery (Lucas et al., 2010).

The neural structures responsible for the mechanisms of liking have only been explored while listening to music. Neuroimaging methods have linked the activation of several cerebral areas to pleasure states when listening to music: the dorsal and ventral striatum, right caudate nucleus, ventral putamen and nucleus accumbens in relation to dopamine release in the reward system (Salimpoor et al., 2011). A functional architecture including anterior cingulate cortex, amygdala, insula, orbitofrontal cortex and nucleus accumbens as areas providing evaluative responses such as the reward value of the stimulus has also been suggested (Salimpoor et al., 2015; Clark et al., 2014). Other structures such as the inferior frontal gyrus (IFG) and the superior temporal gyrus (STG) have been found to be specifically activated during liking/disliking processes (Pereira et al., 2011; Koelsch et al., 2006).

A few recent studies have described the use of EEG in listening situations to investigate the temporal dynamics of liking. Partially discordant results were found in terms of brain rhythms among these studies. When comparing consonant to dissonant music, a positive correlation between the synchronization in the mid-frontal theta band and the liking score emerged, especially towards the end of the excerpts (Sammler et al., 2007). Two other studies (Hadjidimitriou and Hadjileontiadis, 2012, 2013) reported that beta and gamma bands were the best features to separate pleasant from unpleasant music, followed by alpha and theta bands. In a recent study a biomarker of music liking has been defined, relying on brain activations recorded over the left prefrontal cortex and focusing on the functional coupling between high-beta and low-gamma oscillations (Adamos et al., 2016). The discrepancies observed in all EEG studies may be attributed to the interference of the physical processing of sounds that impair the detection of the electrophysiological activity related to liking.

Should liking processes be present during spontaneous auditory imagery, we aimed to investigate whether the spatiotemporal dynamics would be similar to those observed in listening situation. Such a preservation of high-level information processes has already been observed during deliberate auditory imagery. Several studies report similar spatial and temporal processes during the musical perception and the voluntary imagery of a song (Halpern and Zatorre, 1999; Linke and Cusack, 2015; Schaefer et al., 2011; Vlek et al., 2011; Zatorre et al., 1996). The mental perception of a familiar piece of music occurs with a high efficiency, and concerns not only the lyrics but also the tempo (Halpern, 1988; Levitin and Cook, 1996), pitch (Halpern, 1989; Janata and Paroo, 2006) and timbre (Halpern et al., 2004). To study these dynamics, we chose to use high-density EEG, whose spatial resolution allows localizing the neural structures involved in liking processes and whose high temporal resolution is adapted to the time-frequency analysis of these events. Our main hypothesis was based on the presence of similar spatiotemporal mechanisms during imagery to those found during music perception. However, no predicted directions were given for time frequency analysis because of the discordant results obtained in previous EEG studies investigating liking during the perception of music. Specific EEG liking responses were thought to spontaneously emerge, especially in the theta band, with neural activations at least partially similar to those observed during listening situations.

2. Methods

The present study is a complementary analysis of the data from the study of (Gabriel et al., 2016) whose primary objective was to explore the onset of the memory dynamics that lead to the automatic completion of missing lyrics during spontaneous musical imagery.

2.1. Participants

Twenty volunteers (7 male, 13 female, mean age = 32 ± 5 years old) participated in the experiment. All participants were right handed, non-musicians (i.e. had not received any formal musical training or had done so for < 6 years), with no neurologic or psychiatric diseases and normal audition. Written consent was obtained from participants prior to taking part in the study, and all received a monetary compensation of 50€. Research methods were approved by the Committee for the Protection of Human Subjects at the Clinical Investigation Center of Besançon (n°14/458).

2.2. Experimental design

Five familiar songs chosen by the participant and five unfamiliar songs selected by the experimenter among original soundtracks recorded by independent or relatively unknown singers were used. Familiar and unfamiliar songs were paired to belong to the same musical genre, sang by a singer of the same sex. All songs were in participants' native language. Loudness levels of all songs were equalized with the Replay Gain algorithm.

The first EEG recording was performed immediately after song selection. All songs were randomly played over headphones and repeated twice. Each Song was 2 min in length and twenty portions were replaced with two-second silent intervals randomly embedded between 10 and 100 s after song onset and with a distance ranging from 2.5 to 3.5 s between gaps. This resulted in 200 trials within familiar songs and within unknown songs. Silence and song presentation were accomplished using the E-Prime software (Psychology Software Tools Inc. Sharpsburg, PA). Participants were instructed to close their eyes and remain still while passively listened to the song. They were requested to rate each song using a visual analog scale (0: unfamiliar; 10: unfamiliar). The knowledge of each song could thus be checked to ensure that subjects did not know any of the unfamiliar songs. After the EEG recording, subjects were required to rate whether they were mentally completing the gaps during the experiment. As expected, all reported mentally completing the gaps except for unknown songs.

Then there was a two-week learning period during which each subject was inquired to listen regularly to the unfamiliar songs, and to note the frequency of song listened to per day in a self-monitoring logbook. The participants listened to each song for an average of 33.89 times (average = 2.39 listening per day). All subjects listened to each of the five unfamiliar songs the same amount of time.

The second EEG recording was scheduled after the two-week learning period. The recording had exactly the same design as for the first EEG, at the exception of the order of the songs that was randomly selected. At the end of the recording, subjects were requested to rate subjective liking/disliking (from 0: disliked to 10: liked a lot), emotions (from 0: very sad to 10: very joyful) and familiarity (from 0: unfamiliar to 10: very familiar) for each song with a visual analog scale.

2.3. Data acquisition and analysis

In this study, we focused on the liking of newly-learned songs, i.e. unknown songs that were learned over the time course of two weeks, because they were controlled to be listened to the same amount of time. Among the five newly-learned songs listened to by participants, the two more preferred and the two least preferred songs were selected. Only the neural responses from the second EEG were considered.

Subjects were seated in a comfortable chair in a dark quiet testing room. EEG signals were recorded using a 256-channel Geodesic Sensor Net (Electrical Geodesics Inc.; EGI, Eugene, OR). Data were continuously recorded with a high-pass filter at 1 Hz, and a sampling rate at 1000 Hz. Epochs of 2500 ms were extracted from the raw data beginning 600 ms pre-silence onset and ending 1900 ms post-silence onset separately for liked and disliked songs. A total of 80 trials were

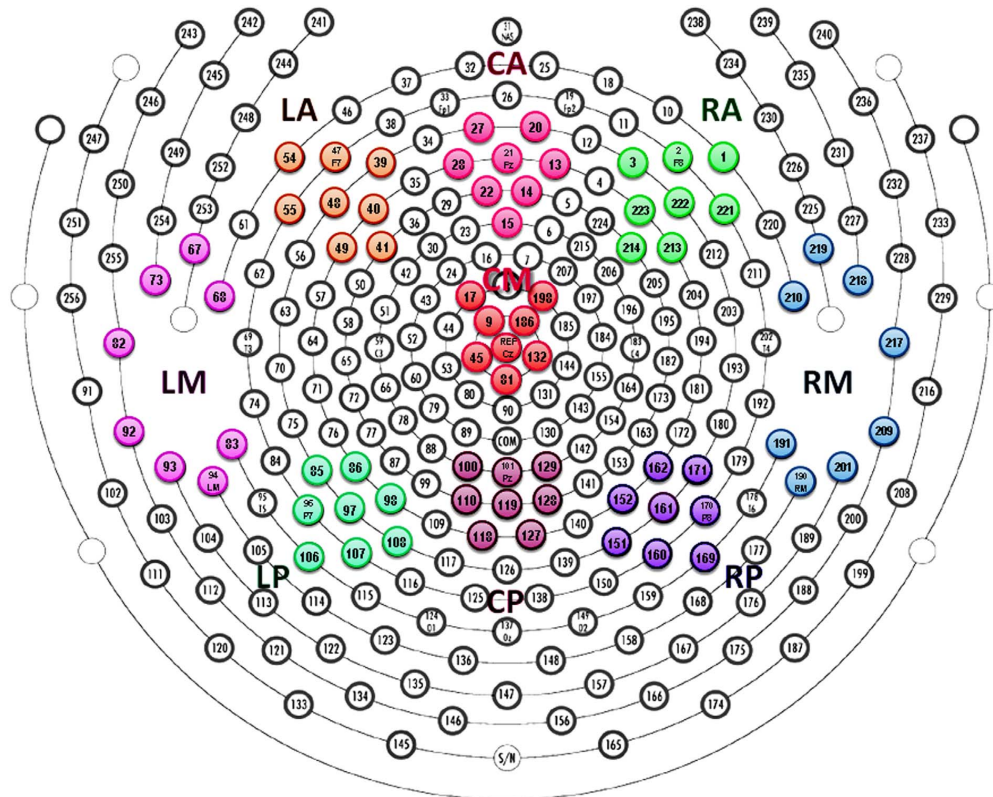


Fig. 1. The nine regions of interest. A: anterior, M: median, P: posterior; L: left, C: central, R: right.

extracted for the more preferred and the least preferred songs. A band-pass filter between 1 and 30 Hz and a notch of 50 Hz were applied to remove unwanted frequency components. A baseline correction was applied for each trial before silence onset (-600 to 0 ms). Periods with visually detectable artifacts were removed from the analysis. Individual channels with artifacts were interpolated using a 3-dimensional spline algorithm (maximum: 7% interpolated electrodes per subject).

Nine Regions of Interest (ROIs) were selected: left anterior, central anterior, right anterior, left median, central median, right median, left posterior, central posterior and right posterior (Fig. 1).

2.3.1. Time-frequency analysis

To evaluate the long-term signal synchronization in the frequency domain during the entire silent period, time-frequency analysis of silent was carried out in the whole time period on each ROI for liked and disliked songs. Spectral power from -50 ms pre-stimulus to 1950 ms post stimulus (in 10 ms steps). The length of the sliding time window was frequency dependent with 3 cycles at the minimum frequency of 3 Hz and 6 cycles at the maximum frequency of 30 Hz. Fifty log-spaced frequencies from 3 Hz to 50 Hz were calculated. We considered as theta the rhythmic activity occurring between 4 and 8 Hz, alpha the activity between 8.5 and 12 Hz, and beta the activity above 12 Hz and below 30 Hz. Time-frequency points were considered significant between liked and disliked songs by a nonparametric permutation-based statistical method (Delorme and Makeig, 2004). A total of 1000 permutations were used at each frequency to assess significance at $p < 0.01$. To increase statistical power, only clusters of significant responses on several frequencies and several time frames were considered as

significant.

2.3.2. Source localization

Based on the results of the time-frequency analysis, source localization was performed on the frequency bands showing significant differences between liked and disliked songs. Source localization was estimated using a distributed linear inverted solution based on a local auto-regressive average (LAURA) model (Grave de Peralta Menendez et al., 2001). Activations obtained by source imaging were compared for each condition and thus identified diverse regional brain activities. Voxelwise two-tailed t -tests were calculated on the mean LAURA source estimates for liked and disliked songs. Differences in the current source density at $p < 0.001$ are reported.

3. Results

3.1. Behavioral analysis

After two weeks of passive learning, the average liking rating was 5.3 ± 2.4 (range 0 – 10). When separating the songs into the two most liked and the most disliked for each subject, a 4 point difference was noted on average (Table 1). Since we selected the most and less liked music for each subject, a highly significant difference was observed between liking scores of liked and disliked songs ($t(19) = 9.55$, $p < 0.001$). The comparison of emotions conveyed by liked and disliked songs revealed no significant difference ($t(19) = 0.81$, $p = 0.4$). Although all songs were listened to for the same amount of time, a significantly increased familiarity for a liked song was observed (t

Table 1

Musical appreciation, familiarity, and emotions for all 5 songs, as well as for the most liked and disliked songs (mean followed by standard deviation). p value shows the difference of scores between liked and disliked songs.

	All songs	Liked songs	Disliked songs	p value
Liking	5.4 ± 2.4	7.3 ± 2.1	3.3 ± 1.7	< 0.000001
Familiarity	7.4 ± 1.6	7.8 ± 1.4	7.0 ± 1.7	< 0.001
Emotions	4.8 ± 2.4	4.5 ± 2.2	5.0 ± 1.6	n.s.

(19) = 4.79; $p < 0.001$).

3.2. Time frequency responses

Time frequency analysis revealed three distinct time periods of significant differences between liked and disliked songs: the first between 400 and 900 ms, the second between 1000 and 1300 ms, and the third between 1400 and 1900 ms (Fig. 2). Differences were mostly significant in the theta and alpha bands, with a desynchronization of theta and a synchronization of alpha for liked songs. No differences were observed above 20 Hz.

Source localization was performed on the three time periods of difference. During the first time period, a significantly stronger theta

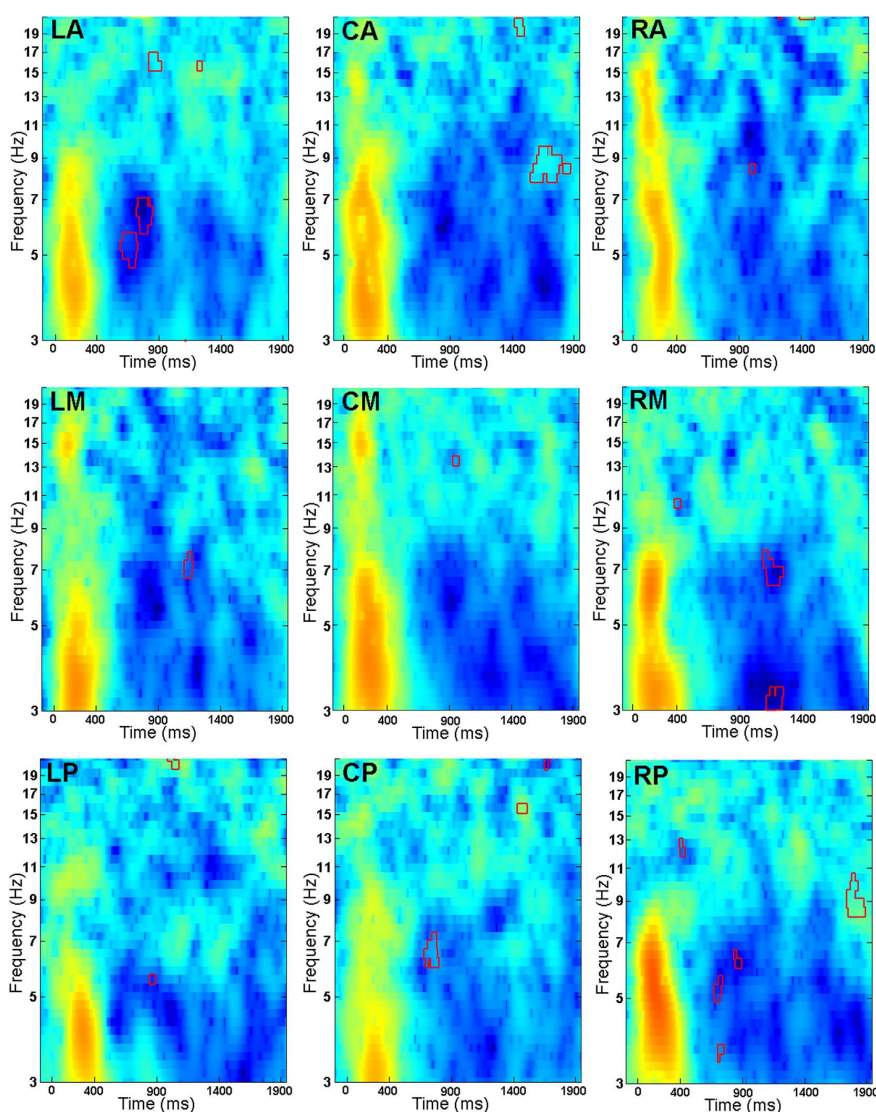


Fig. 2. Time frequency analysis performed over the nine ROIs during silence inserted into liked songs. Neural synchronizations (yellow) and desynchronizations (blue) are represented at different frequency bands. Areas of significant differences ($p < 0.01$) between liked and disliked songs are surrounded in red. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

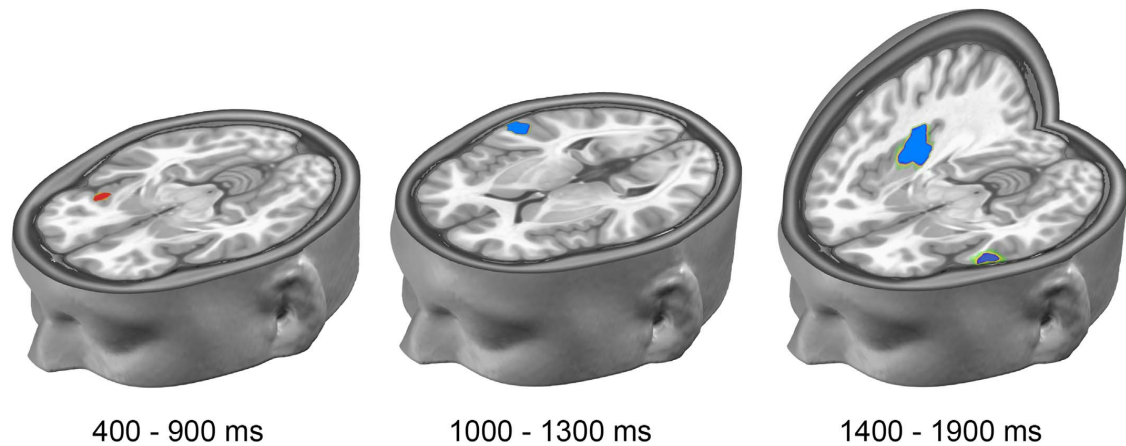


Fig. 3. Brain areas showing significant differences in activation between liked and disliked songs. Left: source of difference in theta activity in the first time period. Middle: source of difference in theta activity in the second time period ($p < 0.001$). Right: source of difference in alpha activity in the last time period ($p < 0.001$). Areas more strongly activated for liked songs are represented in red, and in blue for disliked songs. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

activity was observed in the left anterior and central posterior ROIs (LA and CP in Fig. 2). Statistics performed on source localization could not reveal the source of the differences between both conditions, but the right inferior frontal gyrus (IFG) was differentially activated at a lower significance level of $p < 0.01$ (Fig. 3). In the second time period, a difference in theta activity was observed in the right median ROI (RM in Fig. 2). The main difference originated from a stronger activation of the right superior temporal gyrus (STG) for disliked songs. In the last time period, a difference in alpha activity was observed in the central anterior and the right posterior ROIs (CA and RP in Fig. 2). Source localization showed this difference originated from the left Middle Temporal Gyrus and the right Insula.

4. Discussion

An original aspect of the present study was to explore the neural mechanisms of liking during spontaneous auditory imagery, thus offering the possibility of studying its emergence while removing any acoustical interference. To study music liking in a naturalistic condition, we voluntarily chose songs instead of using artificial excerpts of music and asked the participants to simply listen to these songs for two weeks, as if they were hearing a new song on the radio, for example. Moreover, we preferred using songs as material over using instrumental music because we were interested in the integrated memory representation for text and melody in songs (Crowder et al., 1990; Peretz and Zatorre, 2005). In addition, the songs were not chosen according to the emotions they raised. Although studies on liking usually focused on happy musical extracts, since it is this type of music which most frequently leads to the highest liking scores (Hunter et al., 2008; Vuilleumier and Trost, 2015), it has been shown that sad music can also be highly appreciated by listeners (Brattico et al., 2015). We thus chose the most appreciated and least appreciated songs regardless of their emotional valence. Moreover, since we have averaged over many different songs, based on preference, we thereby canceling out differences related to the stimulus.

Our results show that, although participants were instructed to simply listen to the songs in the experimental tasks, specific neural responses related to aesthetic judgements occurred during spontaneous auditory imagery, which is similar to familiar judgements that have already been noted in identical conditions (Gabriel et al., 2016; Kraemer et al., 2005).

4.1. Electrophysiological correlates of liking

The subjective aesthetic appreciation of a song was reflected by responses in theta and alpha frequency bands in three time periods included in silent intervals. First responses, frontal theta oscillations, have not only been reported in subjects listening to pleasant musical pieces (Sammler et al., 2007), but also in subjects performing a broad range of cognitive tasks requiring attention or short term memory (Burgess and Gruzeliier, 1997; Inanaga, 1998; Smith et al., 1999). Here, it is likely that theta activity observed in the two first time periods reflects both attentional and emotional mechanisms that could be part of the same process since an emotionally significant stimulus automatically attracts attention (Lang et al., 1997). It is, therefore, conceivable that participants were more attentive during silences occurring in pleasant songs, waiting for the music to resume so they could enjoy it again.

Variations in parieto-occipital alpha oscillations observed for liked and disliked songs could be related to auditory attention mechanisms, and it is thus possible that liked music attracted more attention than disliked songs, but this hardly explains why it occurred in such a late period. A suggestion could be that participants were eager for the pleasant songs to start again, and were attentively waiting for the end of silence gaps. However, alternative explanations could explain the increase in alpha oscillations. For example, it has also been suggested that alpha synchronization reflects an active process of inhibition of task-irrelevant areas (Klimesch et al., 2007). As a consequence, variations in alpha oscillations may also be related to increased relaxation consecutive to increased enjoyment and fewer cognitive resources allocated to liked songs (Caldwell and Riby, 2007).

It should be noted that we did not find evidence that oscillations above 20 Hz were involved in music liking, which contradicts some previous studies investigating music liking through EEG signals (Adamos et al., 2016; Hadjidimitriou and Hadjileontiadis, 2012, 2013). In addition to the methodological discrepancies, a major difference between these studies and ours lies in the fact that they measured liking at the same time as music was perceived. In that respect, it is well known that frequencies above 20 Hz fluctuate during musical beat processing (Fujioka et al., 2009; Edagawa and Kawasaki, 2017; Gordon et al., 2011), a parameter especially difficult to control in naturalistic conditions. Studying liking during silence may have reduced the interference between beat processing and liking, even though some studies have shown that tempo was well preserved during auditory

imagery (Halpern, 1988; Levitin and Cook, 1996). Of course, it is also possible that the processing of liking differs during silence and music perception and that aesthetic information is not encoded the same way in both situations. However, the difference in theta oscillations observed in our study, also found when listening to consonant and dissonant music (Sammler et al., 2007), may be more a signature of music liking.

4.2. Neural sources of EEG activity

Interestingly, sources of theta activity observed during silence were found to originate from the IFG and the STG. These results corroborate previous fMRI studies that also reported a stronger activation of the IFG for pleasant music and of the STG for unpleasant music in listening situations (Pereira et al., 2011; Koelsch et al., 2006). The IFG and STG are two structures known to work together to process various aspects of music (Salimpoor et al., 2015). They are activated for musical memory retrieval (Halpern and Zatorre, 1999; Peretz and Zatorre, 2005; Platel et al., 2003) and participate in the formation of a musical image (Sato et al., 2006). The IFG is notably involved in the long-term musical memory retrieval (Altenmüller et al., 2014), which is essential for making accurate predictions about music. The auditory cortex in the superior temporal region is a hub for liked, disliked and favorite music (Wilkins et al., 2014). Our spatiotemporal analysis of subjective liking/disliking during silence revealed that the IFG and the STG do not work simultaneously but successively, increases in liking inducing first a stronger activation of the IFG and then a reduced response of the STG. IFG being known to have an inhibitory role on a number of brain systems and related behaviors (Swann et al., 2009), this structure may inhibit the STG for processing information related to liked songs.

In the last period of silence, the differential activation of the insula suggests that hedonic processes are involved. The insula contributes to the conscious experience of emotion and its activation covaries with subjective feelings, which not only reflect physical stimulus intensity but also cognitive factors such as prediction (Suzuki, 2012).

There were surprisingly few cerebral areas activated for the aesthetic judgement of music in our study, which differs from the large number of cerebral activations observed in fMRI studies (Wilkins et al., 2014; Pereira et al., 2011; Koelsch et al., 2006). Since EEG surface signals are predominantly produced by cortical structures, source localization methods have difficulties to localize deep brain structures. This may explain why structures essential in liking processes such as the limbic system structures and the hippocampus (Wilkins et al., 2014), could not be detected in our study. Moreover, liking scores were not greatly different between liked and disliked songs, only 4 points on average, whereas some studies have been investigating participants' favorite songs (Wilkins et al., 2014) or compared dissonant and consonant music (Sammler et al., 2007). Finally, a possibility is that aesthetic judgement is altered during silence compared to during music perception, leading to a partially activated network of neural activations.

4.3. Relationship between liking, exposure, and familiarity

Since the liking for unfamiliar music is known to increase as a function of exposure time (Schellenberg et al., 2008; Szpunar et al., 2004), we wanted to avoid any confounding effect by assuring that participants listened to liked and disliked songs the same number of times. Despite this control, we found that liked songs were rated as more familiar than disliked songs, confirming that listeners do not only like music that they remember but also remember music that they like (Stalinski and Schellenberg, 2013). This strong relationship between preference and familiarity is supported by source localization. The IFG and STG are two structures involved both in memory processes and in liking. Moreover, emotion-related limbic and paralimbic regions as well as the reward circuitry are known to be significantly more active for

familiar relative to unfamiliar music (Pereira et al., 2011). If musical preference depends on predictive processes based on our knowledge of a song, it is then obvious that familiarity plays an important role. Further research is necessary to clarify this point, although it would be extremely difficult, if not impossible, to measure musical preference without any interference from familiarity and exposure.

5. Conclusion

Results from the present study show that when we like listening to a song, we still like this song during the silences inserted inside. Silences trigger auditory imagery and neural responses specific to liking under the form of oscillations in theta and alpha frequency bands originating from the IFG and STG. These brain structures having previously been shown to be activated when listening to liked music, similar neural networks may be involved in imagery and music. However, all liking processes may not be present because there were less cerebral areas activated for the aesthetic judgement of music during auditory imagery. The experimental design used here nonetheless offer the opportunity to study the emergence of the spatiotemporal dynamics of liking without physical interference of music.

Acknowledgments

This work was supported by a grant from the University Hospital of Besançon (Internal Project named APICHU in 2014: reference no. API/2014/51).

References

- Adamos, D.A., Dimitriadis, S.I., Laskaris, N.A., 2016. Towards the bio-personalization of music recommendation systems: a single-sensor EEG biomarker of subjective music preference. *Inf. Sci.* 343-344, 94-108.
- Altenmüller, E., Sigel, S., Mohammadi, B., Samii, A., Munte, T.F., 2014. Play it again, Sam: brain correlates of emotional music recognition. *Front. Psychol.* 5, 114.
- Blood, A.J., Zatorre, R.J., 2001. Intensely pleasurable responses to music correlate with activity in brain regions implicated in reward and emotion. *Proc. Natl. Acad. Sci. U. S. A.* 98 (20), 11818-11823.
- Brattico, E., Bogert, B., Alluri, V., Tervaniemi, M., Eerola, T., Jacobsen, T., 2015. It's sad but I like it: the neural dissociation between musical emotions and liking in experts and laypersons. *Front. Hum. Neurosci.* 9, 676.
- Burgess, A.P., Gruzelier, J.H., 1997. Short duration synchronization of human theta rhythm during recognition memory. *Neuroreport* 8 (4), 1039-1042.
- Caldwell, G.N., Riby, L.M., 2007. The effects of music exposure and own genre preference on conscious and unconscious cognitive processes: a pilot ERP study. *Conscious. Cogn.* 16 (4), 992-996.
- Clark, C.N., Downey, L.E., Warren, J.D., 2014. Music biology: all this useful beauty. *Curr. Biol.* 24 (6), R234-237.
- Crowder, R.G., Serafine, M.L., Repp, B., 1990. Physical interaction and association by contiguity in memory for the words and melodies of songs. *Mem. Cogn.* 18 (5), 469-476.
- 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 (1), 9-21.
- Edagawa, K., Kawasaki, M., 2017. Beta phase synchronization in the frontal-temporal-cerebellar network during auditory-to-motor rhythm learning. *Sci. Rep.* 7, 42721.
- Fujioka, T., Trainor, L.J., Large, E.W., Ross, B., 2009. Beta and gamma rhythms in human auditory cortex during musical beat processing. *Ann. N. Y. Acad. Sci.* 1169, 89-92.
- Gabriel, D., Wong, T.C., Nicolier, M., Giustiniani, J., Mignot, C., Noiret, N., et al., 2016. Don't forget the lyrics! Spatiotemporal dynamics of neural mechanisms spontaneously evoked by gaps of silence in familiar and newly learned songs. *Neurobiol. Learn. Mem.* 132, 18-28.
- Gordon, R.L., Magne, C.L., Large, E.W., 2011. EEG correlates of song prosody: a new look at the relationship between linguistic and musical rhythm. *Front. Psychol.* 2, 352.
- Grave de Peralta Menendez, R., Gonzalez Andino, S., Lantz, G., Michel, C.M., Landis, T., 2001. Noninvasive localization of electromagnetic epileptic activity. I. Method descriptions and simulations. *Brain Topogr.* 14 (2), 131-137.
- Guity, S., 1947. Toutes réflexions faites. In: Cinquante ans d'occupations. Presses de la cité, coll "Omnibus".
- Hadjilimitriou, S.K., Hadjileontiadis, L.J., 2012. Toward an EEG-based recognition of music liking using time-frequency analysis. *IEEE Trans. Biomed. Eng.* 59 (12), 3498-3510.
- Hadjilimitriou, S.K., Hadjileontiadis, L.J., 2013. EEG-based classification of music appraisal responses using time-frequency analysis and familiarity ratings. *IEEE Trans. Affect. Comput.* 4 (2), 161-172.
- Halpern, A.R., 1988. Mental scanning in auditory imagery for songs. *J. Exp. Psychol.*

- Learn. Mem. Cogn. 14 (3), 434–443.
- Halpern, A.R., 1989. Memory for the absolute pitch of familiar songs. *Mem. Cogn.* 17 (5), 572–581.
- Halpern, A.R., Zatorre, R.J., 1999. When that tune runs through your head: a PET investigation of auditory imagery for familiar melodies. *Cereb. Cortex* 9 (7), 697–704.
- Halpern, A.R., Zatorre, R.J., Bouffard, M., Johnson, J.A., 2004. Behavioral and neural correlates of perceived and imagined musical timbre. *Neuropsychologia* 42 (9), 1281–1292.
- Hunter, P.G., Schellenberg, E.G., Schinmack, U., 2008. Mixed affective responses to music with conflicting cues. *Cognit. Emot.* 22 (2), 327–352.
- Inanaga, K., 1998. Frontal midline theta rhythm and mental activity. *Psychiatry Clin. Neurosci.* 52 (6), 555–566.
- Janata, P., Paroo, K., 2006. Acuity of auditory images in pitch and time. *Percept. Psychophys.* 68 (5), 829–844.
- Klimesch, W., Sauseng, P., Hanslmayr, S., Gruber, W., Freunberger, R., 2007. Event-related phase reorganization may explain evoked neural dynamics. *Neurosci. Biobehav. Rev.* 31 (7), 1003–1016.
- Koelsch, S., Fritz, T., DY, V.C., Muller, K., Friederici, A.D., 2006. Investigating emotion with music: an fMRI study. *Hum. Brain Mapp.* 27 (3), 239–250.
- Kraemer, D.J., Macrae, C.N., Green, A.E., Kelley, W.M., 2005. Musical imagery: sound of silence activates auditory cortex. *Nature* 434 (7030), 158.
- Lang, P.J., Bradley, M.M., Cuthbert, B.N., 1997. Motivated attention: affect, activation, and action. In: Lang, P.J., Simons, R.F., Balaban, M.T. (Eds.), *Attention and Orienting: Sensory and Motivational Processes*. Lawrence Erlbaum Associates, Hillsdale, NJ, pp. 97–135.
- Levitin, D.J., Cook, P.R., 1996. Memory for musical tempo: additional evidence that auditory memory is absolute. *Percept. Psychophys.* 58 (6), 927–935.
- Linke, A.C., Cusack, R., 2015. Flexible information coding in human auditory cortex during perception, imagery, and STM of complex sounds. *J. Cogn. Neurosci.* 27 (7), 1322–1333.
- Lucas, B.J., Schubert, E., Halpern, A.R., 2010. Perception of emotion in sounded and imagined music. *Music. Percept.* 27 (5), 399–412.
- Pereira, C.S., Teixeira, J., Figueiredo, P., Xavier, J., Castro, S.L., Brattico, E., 2011. Music and emotions in the brain: familiarity matters. *PLoS One* 6 (11), e27241.
- Peretz, I., Zatorre, R.J., 2005. Brain organization for music processing. *Annu. Rev. Psychol.* 56, 89–114.
- Platel, H., Baron, J.C., Desgranges, B., Bernard, F., Eustache, F., 2003. Semantic and episodic memory of music are subserved by distinct neural networks. *NeuroImage* 20 (1), 244–256.
- Salimpoor, V.N., Benovoy, M., Larcher, K., Dagher, A., Zatorre, R.J., 2011. Anatomically distinct dopamine release during anticipation and experience of peak emotion to music. *Nat. Neurosci.* 14 (2), 257–262.
- Salimpoor, V.N., Zald, D.H., Zatorre, R.J., Dagher, A., McIntosh, A.R., 2015. Predictions and the brain: how musical sounds become rewarding. *Trends Cogn. Sci.* 19 (2), 86–91.
- Sammler, D., Grigutsch, M., Fritz, T., Koelsch, S., 2007. Music and emotion: electrophysiological correlates of the processing of pleasant and unpleasant music. *Psychophysiology* 44 (2), 293–304.
- Satoh, M., Takeda, K., Nagata, K., Shimosegawa, E., Kuzuhara, S., 2006. Positron-emission tomography of brain regions activated by recognition of familiar music. *AJNR Am. J. Neuroradiol.* 27 (5), 1101–1106.
- Schaefer, R.S., Vlek, R.J., Desain, P., 2011. Music perception and imagery in EEG: alpha band effects of task and stimulus. *Int. J. Psychophysiol.* 82 (3), 254–259.
- Schellenberg, E.G., Peretz, I., Veillard, S., 2008. Liking for happy- and sad-sounding music: effects of exposure. *Cognit. Emot.* 22 (2), 218–237. <http://dx.doi.org/10.1080/02699930701350753>.
- Smith, M.E., McEvoy, L.K., Gevins, A., 1999. Neurophysiological indices of strategy development and skill acquisition. *Brain Res. Cogn. Brain Res.* 7 (3), 389–404.
- Stalinski, S.M., Schellenberg, E.G., 2013. Listeners remember music they like. *J. Exp. Psychol. Learn. Mem. Cogn.* 39 (3), 700–716.
- Suzuki, A., 2012. Emotional functions of the insula. *Brain Nerve* 64 (10), 1103–1112.
- Swann, N., Tandon, N., Canolty, R., Ellmore, T.M., McEvoy, L.K., Dreyer, S., et al., 2009. Intracranial EEG reveals a time- and frequency-specific role for the right inferior frontal gyrus and primary motor cortex in stopping initiated responses. *J. Neurosci.* 29 (40), 12675–12685.
- Szpunar, K.K., Schellenberg, E.G., Pliner, P., 2004. Liking and memory for musical stimuli as a function of exposure. *J. Exp. Psychol. Learn. Mem. Cogn.* 30 (2), 370–381.
- Vlek, R.J., Schaefer, R.S., Gielen, C.C., Farquhar, J.D., Desain, P., 2011. Shared mechanisms in perception and imagery of auditory accents. *Clin. Neurophysiol.* 122 (8), 1526–1532.
- Vuilleumier, P., Trost, W., 2015. Music and emotions: from enchantment to entrainment. *Ann. N. Y. Acad. Sci.* 1337, 212–222.
- Wilkins, R.W., Hodges, D.A., Laurienti, P.J., Steen, M., Burdette, J.H., 2014. Network science and the effects of music preference on functional brain connectivity: from Beethoven to Eminem. *Sci. Rep.* 4, 6130.
- Zatorre, R.J., Halpern, A.R., Perry, D.W., Meyer, E., Evans, A.C., 1996. Hearing in the mind's ear: a PET investigation of musical imagery and perception. *J. Cogn. Neurosci.* 8 (1), 29–46.

ANNEXE C : AUTRES ARTICLES PUBLIÉS PENDANT LA THÈSE



The Relevance of a Conductor Competition for the Study of Emotional Synchronization Within and Between Groups in a Natural Musical Setting

Thibault Chabin^{1*}, Grégory Tio^{1,2}, Alexandre Comte^{1,2,3}, Coralie Joucla¹, Damien Gabriel^{1,2,3} and Lionel Pazart^{1,2}

¹ Laboratoire de Neurosciences Intégratives et Cliniques EA481, Université Bourgogne Franche-Comté, Besançon, France,

² Centre Hospitalier Universitaire de Besançon, Centre d'Investigation Clinique INSERM CIC 1431, Besançon, France,

³ Plateforme de Neuroimagerie Fonctionnelle et Neurostimulation - Neuraxess, Besançon, France

OPEN ACCESS

Edited by:

Rafael Ramirez,
Pompeu Fabra University, Spain

Reviewed by:

Gualtiero Volpe,
University of Genoa, Italy
George Waddell,
Royal College of Music,
United Kingdom

*Correspondence:

Thibault Chabin
tchabin@edu.univ-fcomte.fr;
chabinthibault@yahoo.fr

Specialty section:

This article was submitted to
Performance Science,
a section of the journal
Frontiers in Psychology

Received: 02 May 2019

Accepted: 12 December 2019

Published: 17 January 2020

Citation:

Chabin T, Tio G, Comte A,
Joucla C, Gabriel D and Pazart L
(2020) The Relevance of a Conductor
Competition for the Study
of Emotional Synchronization Within
and Between Groups in a Natural
Musical Setting.
Front. Psychol. 10:2954.
doi: 10.3389/fpsyg.2019.02954

Group emotional dynamics are a central concern in the study of human interaction and communication. To study group emotions, the social context of a musical event in natural conditions may overcome several limits of laboratory experiments and could provide a suitable framework. This study aimed to evaluate if cultural events such as a conductor competition could welcome scientific research for the study of group emotional sharing. We led an observational study, which suggests that in this particular context, public, musicians and jury would agree to participate and to wear neurophysiological and physiological devices to monitor their emotional state during the competition. Self-administrated scales showed that, in the context of a musical competition, members of the public felt strong musical emotions such as music chills. Our results suggest that such a specific competition design is a suitable experimental model to lead an experiment under ecological conditions to effectively investigate collective emotional synchronization. In the future, with the implementation of an acquisition system recording synchronous neurophysiological data for a large group of participants, we may be able to highlight mechanisms involved in emotional synchronization in a natural musical setting.

Keywords: emotional synchronization, EEG, hyperscanning methods, music, conductor competition

INTRODUCTION

Most people listen to music for its emotional effect (Nieminen et al., 2012) and research has shown the pleasurable effect of music (Grewe et al., 2007; Salimpoor et al., 2009, 2011). The social connection with other people and being connected to the performers are the main factors that motivate people to attend live performances and enjoy them (Burland and Pitts, 2014; Leante, 2016; Swarbrick et al., 2019). Music is a form of cultural communication (McCormick, 2009). According to the model of the Share Affective Motion Experience (Molnar-Szakacs et al., 2011), which claims music as “a series of intentional, expressive motor acts, recruiting similar neural networks in both

agent and listener” (Overy, 2012: 66), music can enhance emotions and create a sense of social bonding (Overy, 2012; Chanda and Levitin, 2013).

Considering research that has investigated the effect of music listening in groups, some points out that context (laboratory or natural conditions) and the social environment (being alone or with other people) could greatly influence the experience of listening to music (Lamont, 2009; Liljeström et al., 2013). Being in the context of a concert in the presence of other people can enhance the listening experience by eliciting at least a rhythmic synchronization when people move together (Desmet et al., 2010). Live conditions provide listeners with the strongest musical experiences (Lamont, 2011), with a greater effect on physiological variables as heart rate (Shoda et al., 2016). However, contradictory results have emerged from several studies. On the one hand, more “pleasure enjoyment” was experienced during music listening with a close friend or a partner versus alone (Juslin et al., 2008), and people who listened to music alone were less likely to experience intense emotions compared to people listening with a close friend or a partner (Liljeström et al., 2013). On the other hand, an intra-individual variability has been demonstrated. Some people were more likely to experience chills when listening to music alone, whereas others experienced more when they were with others in daily life situations (Nusbaum et al., 2013). Furthermore, in a laboratory setting, the social context negatively affected the emotional experience: the experience of chills, and thus emotional arousal, was lower in groups (Egermann et al., 2011). The social context, particularly in this laboratory setting experiment, influenced participants’ concentration, altering the emotional arousal (Egermann et al., 2011; Koehler and Broughton, 2017). However, these studies in laboratory or natural contexts with dyads or groups involved major differences in implementation.

Ecological conditions may be of great importance when considering the precise mechanisms of human interactions (Acquadro et al., 2016). An ecological context should clarify and help generalize previous laboratory findings (Egermann et al., 2013), surpassing dyadic interactions (Tarr et al., 2014) and going beyond some of the experimental laboratory limits in the study of group emotions. Collective music performances are a suitable tool through which to study interactions in ecological conditions (D’Ausilio et al., 2015). Thus, to study interpersonal synchronization and the mechanisms involved in social interactions, it is possible to simulate research protocols that are as realistic as possible whilst maintaining ecological validity. Ecological experiments in the musical context have already measured head movement synchronization by simulating a live concert in a living lab (Swarbrick et al., 2019), or measured the heart rate of a few people among a large audience listening to a piano performance in an auditorium (Shoda et al., 2016). Otherwise, when natural settings are not reproducible in ecological conditions, it is possible to benefit from the specific design arising from a natural situation.

By definition natural experiments are studies conducted in a natural everyday life context seeking to measure variations in natural processing without the manipulation of variables (Salkind, 2010). To set up a study in natural conditions,

the setting should be stable over time and should include natural exogenous variations without any interventions from the experimenters (Meyer, 1995; Craig et al., 2012). The drawback of such natural conditions is that the selection of participants, the adherence of subjects with the research, the implementation of subjective measures, the adaptability of the paradigm to environmental conditions and so on need to be evaluated precisely before implementing a study. In this respect, we hypothesized that the specific context of a conductor competition would bring an advantageous framework to study the spread of emotions and emotional synchronization within and between groups (music performers and listeners) in natural conditions. In this context, both cerebral activity, in terms of natural activation of the autonomic nervous system, and subjective feeling could be recorded among the different participants (orchestra performers, jury, candidates, and public) during the competition.

In the specific context of the conductors’ competition, the same orchestra performing the same musical pieces in the same concert hall in front of the same public offers favorable live setting conditions. The spread of emotions can be assessed within and between groups of performers and listeners. Contrary to laboratory settings, the musical material is ecological in this context. Direct reports of emotional responses from participants and neurophysiological recordings should give objective and subjective measures for the evaluation of groups’ emotional synchronization. In this context, any intervention of the experimenter would influence the responses. Furthermore, the natural response of the participants should be controlled at a physiological level. Music chills can be defined as a strong transient emotional phenomenon triggered by specific musical features and involving reactions of the brain and body in the form of goose bumps and thrills, and which can provoke intense pleasure (Goldstein, 1980; Salimpoor et al., 2011). Chills present the advantage of being identified both by physiological measures and directly reported by the participant (Grewe et al., 2009a,b, 2011). Blood and Zatorre (2001) suggest that pleasantness and emotional intensity must increase to a certain level before the occurrence of a chill. Musical chills that are the climax of an increase in emotion (Grewe et al., 2009a) should be a reliable indicator of a strong emotional state. They can be generated in a social situation such as a concert, movie theater, or other gathering where people are involved in the same action and focused on the same target (Schoeller, 2015). In that respect, music chills could indirectly be an indicator of common emotions between several participants when chills are experienced at the same time. Using this observational approach should allow experimenters to record a set of data on different interacting populations, wherein synchronization can be evaluated in a natural way, without no intervention required, throughout the competition.

Volpe et al. (2016) insisted on acquiring large data samples in valid ecological conditions using multimodal tools to assess synchronization. Measuring subjective feelings and peripheral arousal during live concerts provides a great opportunity to study the influence of other people on emotional responses to music (Egermann et al., 2013). The new ambulatory monitoring techniques of physiological recording could highlight numerous

as yet unknown fundamental mechanisms and datasets related to information in the field of emotion research, especially in studies in ecological conditions (Wilhelm and Grossman, 2010; Shoval et al., 2018). Some musical research has already been coordinated in both passive and active conditions. A number of studies have shown the effect of listening to music in terms of the emotional response using physiological recording, considering both cardiac and electrodermal activities (Craig, 2005; Grewe et al., 2011; Mori and Iwanaga, 2014; Colver and El-Alayli, 2016). Moreover, the development of electroencephalographic (EEG) hyperscanning methods allows researchers to record EEG data on several participants simultaneously, providing an appropriate tool for the study of brain processes involved in social interactions (Dumas et al., 2010; Acquadro et al., 2016; Toppi et al., 2016; Balconi and Vanutelli, 2017).

In line with this approach, we led an observational study that is a first step in the implementation of natural experiments using neurophysiological sensors in natural conditions. The overall goal was to assess the appropriateness of using a conductor competition to study emotional synchronization within and between groups (music performers and listeners) in natural conditions. Three dimensions were assessed during this observational study: the acceptability for listeners (public, jury) and performers (musicians, conductors) of wearing different sensors during the competition; the emotional sensitivity of the public in a such context; and the suitability of competition settings to replicate an experimental design in natural conditions.

MATERIALS AND METHODS

Ethics Statement

The research met the local ethical rules provided by French law on non-invasive protocols involving healthy participants, and was classified as a psychological observational study outside the Jardé law (Article R1121-1 of the French Law Code of Public Health amended by decree n° 2017-884 of May 2017). It was submitted to the Ethics Committee CPP Est II, which exempted the study from the full ethics review process. Each participant was informed of the observational nature of the study. Participants freely filled out questionnaires and gave their implicit consent by returning the questionnaire. The questionnaires were anonymous, and no participants can be identified. Participants who were interviewed gave their explicit oral consent before answering the investigator's questions.

Presentation of the International Competition for Young Conductors

The 55th Besançon International Competition for Young Conductors in France is a musical event based on the musical direction of various styles and various ensembles (two pianos, opera, symphonic orchestra, etc.). From the pre-selection to the final phase, every candidate has to perform the musical direction of the same musical extracts. Our study concerns only the final phases. The final rounds of the competition lasted several days and were composed of two sessions per day, the first during the afternoon and the second in the evening. We

named data collection sessions according to the day (Monday: Mon; Tuesday: Tue; Wednesday: Wed) and according to the time of day: afternoon session [1] or evening session [2]. We collected data in five sessions: during the 1/8 elimination round (Mon1 and Mon2), the 1/4 final round (Tue1 and Tue2), and the first semi-final round (Wed1). To identify the candidates engaged in the competition, letters from A to T were assigned to each of the 20 conductors.

The 1/8 elimination round took place at the Grand Kursaal on Monday, 11 September 2017. Ten candidates during the first session (Mon1) and ten candidates during the second session (Mon2) conducted the Orchestre Victor Hugo de Besançon Franche-Comté using three extracts of approximately 10 min from the musical piece "Appalachian's Spring Orchestral Suite" by Aaron Copland. The 1/4 elimination round took place at the Ledoux Theater on Tuesday, 12 September. Six candidates during the first session (Tue1) and six candidates during the second (Tue2) conducted the Orchestre Victor Hugo de Besançon Franche-Comté using three extracts of approximately 20 min from "Carmen" by Georges Bizet. The first semi-final round took place at the Grand Kursaal and included six candidates, all of whom participated in both sessions. First, they conducted the Orchestre National de Lyon using two extracts lasting approximately 25 to 30 min from the 'Concerto No.2 for Piano and Orchestra in F minor, Op.21' by Frédéric Chopin (Soloist: Philippe Cassard) (Wed1). For further details on the last rounds of the competition, see **Supplementary Data 1**.

Location

The final rounds took place in two concert halls in Besançon: the Grand Kursaal and the Ledoux Theater. The Grand Kursaal is a 480 m² concert hall containing two balconies and a great stage; it is able to accommodate approximately 800 people. The Ledoux Theater is a classic theater composed of a stage and a pit for the orchestra; it is able to accommodate 1,000 persons on three levels.

Hypothesis and Objectives

We hypothesized that the specific context of this conductor competition would provide an advantageous framework to measure the emotional synchronization within and between groups (music performers and listeners) in natural conditions using neurophysiological devices.

To test this hypothesis, three requirements must be addressed:

1. *Acceptability of further study*: Will participants agree to wear neurophysiological sensors during the competition? (Feasibility for recruitment).
2. *Public's emotional sensitivity and public profile*: Will public members be sensitive to musical emotions and to music chills? Will they be representatives of the general public of classical concert? (External validity).
3. *Suitability of competition setting*: Will the extracts from each music piece be performed by each conductor for a minimal duration of 1 min/30 s? (Reproducibility of similar conditions) Could some confounding factors (e.g., types of music piece, orchestra, architecture of the concert hall) be managed to avoid measurement bias? Will a significant

number of conductors produce music chills in the public? (At least 30% should experience chills).

Study Design

In this observational study, different activities were conducted to answer our three main requirements.

1. Acceptability for a Further Study

We conducted face to face interviews to assess the willingness of some public members, jury members, musicians, and conductors to wear sensors (EEG headsets and wristbands recording physiological data) during the conductor competition.

2. Public's Emotional Sensitivity and Public Profile

We distributed self-reported questionnaires to evaluate personality profiles and public emotions in the specific context of the competition, involving a lot of repetitions of the same short extract of music pieces.

3. Suitability of Competition Setting

We did observations of the competition setting to ensure the right reproducibility of each candidate's performance for further statistical analysis.

Population

To assess the acceptability of the research protocol, the different groups were jury members ($n = 7$), public members (around 300 to 400 per session), musicians ($n = 45$), and conductors ($n = 20$) engaged in the competition. See **Table 1** for demographic information about the participants involved. Interviewees were selected based on their availability during the breaks (before, during and after the performance). To test the public's musical emotional sensitivity, the only eligibility criterion considered for the distribution of all questionnaires and for collecting opinions was to be aged over 18 years. Participation was on a voluntary basis.

Variables, Data Sources and Measurement

Acceptability of Further Study

A sample of the four groups (orchestra performers, jury, candidates, and public) was questioned orally about the acceptability of both the type of mobile EEG headset and the tools that could be used in such conditions (three different mobile devices were presented in pictures), along with whether they would agree to participate in a similar study about emotions

in the next edition of the competition. The interviews were conducted individually by TC and LP, mainly during the intermissions, detailing our aim, the setting for the study and the different requirements involved in participation. Because of competition conditions, interviews were conducted directly in the concert rooms and personal information about age and gender were not collected during oral interviews. The only criterion was to be aged over 18 years, with no further discrimination. The mobile tools presented were EEG wireless headsets – namely Emotiv Insight, Emotiv EPOC+ and Imec's wireless EEG headset – and connected wristbands such as Jawbone UP3 and Fitbit. Oral interviews were conducted individually with candidates, musicians and members of the jury during intermissions. The interviews were conducted with the musicians and with the conductors during the 1/8 final round (Mon1, Mon2). They were conducted with members of the public during the five data collection sessions (Mon1, Mon2, Tue1, Tue2, and Wed1).

Public's Emotional Sensitivity: A Public Profile

Self-reported questionnaires were used to ascertain the public profile. Each questionnaire gave only the age and sex of the responder. Questionnaires were distributed at each entrance to the concert hall before the performance. Participants freely filled out and returned the questionnaire at the end of the session.

Many psychology studies have pointed out the positive link between the personality trait of openness to experience and sensitivity to aesthetic/music chills (Grewe et al., 2007; McCrae, 2007; Nusbaum and Silvia, 2011a; Silvia and Nusbaum, 2011; Sumpf et al., 2015; Colver and El-Alayli, 2016). We chose to assess the emotional sensitivity of responders and their sensitivity to music chills with a French adaptation of the Aesthetic Experiences Scale in Music or AES-M (Nusbaum and Silvia, 2011b; Sachs et al., 2016). Finally, we used an adaptation of the AES-M by candidate (emotional scale by candidate, ESC) to make comparisons between each candidate and each performance. Since musical emotional sensitivity could be linked with the personality profile, we also used the French adaptation of the Ten Item Personality Inventory (TIPI), extracted from the Big Five personality domains (BFI) (Gosling et al., 2003). Using the personality profile increased the internal validity of our data. Finally, the classification of the French Orchestra Association (AFO, Association Française des Orchestres, 2014) was used to compare public from the conductor competition with public from classical concert. All these scales are presented in brief below and in **Supplementary Appendices 1–3**.

Ten Item Personality Inventory

The TIPI (see **Supplementary Appendix 1**) is composed of ten items, five of which are opposed to the other five. This allows the researchers to calculate five personality trait scores: extraversion, agreeableness, conscientiousness, emotional stability, and openness to experience. Each personality trait score is calculated by averaging the score of one item with the reverse score of its opposite item on a seven-point Likert scale. The reverse score corresponds to the reverse rating of the score (for example, if the score is one, the reverse score will be seven;

TABLE 1 | Public sample by session.

Session	Questionnaires distributed	Questionnaires returned (%)	Males/ Females (%)	Mean age (SD)
Mon1	90	45 (50)	34/66	60.9 (15)
Mon2	71	23 (32.3)	24/76	54.5 (16.4)
Tue1	120	21 (17.5)	19/81	67.5 (8.4)
Tue2	150	39 (26)	25/75	50.7 (16.6)
Wed1	150	38 (25.3)	30/70	63.5 (14)

if the score is two, the reverse score will be six, etc.). The item “extraverted, enthusiastic” is opposed to the item “reserved, quiet,” giving a score for the extraversion personality trait. The item ‘critical, quarrelsome’ is opposed to the item “sympathetic, warm,” giving a score for the agreeableness personality trait. The item “dependable, self-disciplined” is opposed to the item “disorganized, careless,” giving a score for the conscientiousness personality trait. The item “anxious, easily upset” is opposed to the item “calm, emotionally stable,” giving a score for the emotional stability personality trait. Finally, the item ‘open to new experiences, complex’ is opposed to the item “conventional, uncreative,” giving a score for the openness to experience personality trait.

Aesthetic experience scale in music

Based on Nusbaum and Silvia’s previous work (Nusbaum and Silvia, 2011b) (see **Supplementary Appendix 2**), four factors were calculated for the AES-M from ten items: the chill factor, the touched factor, the absorption factor and the global factor. The chill factor is composed of the items “Feel chills down your spine,” “Get goose bumps” and “Feel like your hair is standing on end,” giving scores out of three. The absorption factor is composed of the items “Feel absorbed and immersed,” “Completely lose track of time,” “Feel like you’re somewhere else,” “Feel detached from your surroundings” and “Feel a sense of awe and wonder,” giving scores out of five. The touched factor is composed of the items “Feel touched” and “Feel like crying,” giving scores out of two. The score for a factor indicates how many items of this factor had a response greater than or equal to four on a seven-point Likert scales (corresponding to “sometimes”). The global factor score out of ten was calculated for each participant by adding the three factor scores.

Emotional scale by candidate

This scale (see **Supplementary Appendix 3**) tests whether people can feel different emotions or musical chills according to the conductor. The scale by candidate is composed of the items extracted from the AES-M that participants checked off for each candidate. This allows the researchers to calculate four factors similar to those of the AES-M. The chill factor gives scores out of three, the absorption factor gives scores out of five, and the touched factor gives scores out of two. A global factor score out of ten was calculated for each participant by adding the three factor scores. The absence of a response was considered as no feeling for the item by a candidate.

French Orchestra Association (AFO) Classification

During the semi-final, people were asked to define their auditor profile by choosing between the profiles extracted from a classification of the French Orchestra Association (Association Française des Orchestres, 2014). These categories were “classical music lover”: first, interested in classical music and then, sharing a moment in a beautiful place; “curious music lover”: focused on a musical experience, interested in discovering new musical works and searching for new emotions aroused by musical work; “sociable”: concert experience is a pleasurable moment shared with friends and family in the musical emotion register; “occasional layman”: infrequent musical and concert practice, came with family or friend and enjoyed the spectacular dimension of the event; “distanced amateur”: concerts are an uncommon opportunity and not a usual practice.

Course of the observational study

The schedule of different activities is presented in **Table 2**.

Suitability of Competition Settings

Finally, direct observations of the concert hall and of the competition process were performed by noting the timing of the performance and by analyzing the musical work to determine whether the competition was adapted to neurophysiological recordings.

Statistical Analysis

Statistical analysis only covers self-reported questionnaires to describe the public profile.

A Shapiro–Wilk test was used to check that the data as a whole followed a normal distribution. Descriptive statistics were analyzed for the age, gender and size of the sample to determine the characteristics of the public who responded to the questionnaires during each session (presented in **Table 3**). For all statistical tests, the p -value was fixed with $\alpha = 0.05$ and analysis involved a Bonferroni’s correction based on the multiplicity of the tests. The comparison between sessions for TIPI scores was analyzed for all participants using analysis of variance (ANOVA) for extraversion and emotion traits, and a Kruskal–Wallis test for agreeableness, consciousness and openness to experience traits (correction of the p -value threshold to $p = 0.01$).

Emotional Sensitivity According to Candidates

A non-parametric repeated measures ANOVA GLIMMIX procedure under SAS was applied for each ESC factor for Tue2

TABLE 2 | Schedule of activities.

Activities	Mon1	Mon 2	Tue 1	Tue 2	Wed 1
	First 1/8 elimination round (10 conductors)	Second 1/8 elimination round (10 conductors)	First 1/4 round (6 conductors)	Second 1/4 round (6 conductors)	First semi-final (6 conductors)
Interviews	X	X	X	X	
AES-M	X	X	X		
ESC				X	X
TIPI	X	X	X	X	
AFO					X
Direct observation of setting	X	X	X	X	X

TABLE 3 | Acceptability of further study.

Categories	Size of the population	Size of the sample	Agreement, number (%)
Jury	7	5	5 (100)
Public	400	36	34 (94)
Musicians	45	15	15 (100)
Conductors	20	13	9 (70)
Total		69	63 (91)

and Wed1 conductors (correction of the p -value threshold to $p = 0.0125$) to compare the ratings for each of the conductors (random effect), followed by a *post hoc* contrast analysis with a p -value correction (p -value threshold: 0.0033). Effect size were calculated with Kendall coefficient of concordance (W). For the ESC, a descriptive analysis of the chills felt (indicating the number of people who reported at least one item of the chill factor) and the number of people who reported chills for other candidates are presented in **Table 4**.

AFO Classification

Chi² tests were applied to compare the results of the AFO classification in our study and in the AFO survey. Effect size were calculated with Cramer's V .

RESULTS

Acceptability of Further Study

Sixty-nine individuals were questioned orally during the whole competition. **Table 3** shows the sample size of the different groups (members of the jury, members of the public, members of the orchestra and conductors engaged in the competition) that agreed to wear sensors for a future study.

Among the 20 conductors, 13 were interviewed, of whom eight gave their agreement; one more agreed as long as devices were worn by all candidates. Among the four conductors who refused to wear sensors, three of them indicated the possible discomfort of the EEG headset as they already wear a microphone, and the fourth cited the distraction the sensor could cause. All of them agreed to participate in the study if the research protocol was included in the competition without wearing sensors. In summary, 91% of people interviewed agreed to wear sensors to participate in a future study during the conductor competition.

Public's Emotional Sensitivity and Public Profile

Five hundred and eighty one questionnaires were distributed during five sessions of the competition, of which 166 were returned, giving a response rate of 28.6%. See **Table 1** for details.

Regarding the TIPI scales, 128 questionnaires were returned. There were no significant differences in sessions for any items of the TIPI.

Regarding the AES-M, 89 questionnaires were returned. Descriptive results indicate that the factor "touched" was rated

greater than or equal to four for 85 participants during the competition. This suggests that 95.6% of the participants indicated that they felt "touched" during the competition. In the same way, 70 participants (78.7%) indicated being "absorbed" and 29 (32.6%) indicated feeling chills. Finally, the global factor indicated that there were 26 participants (29.2%) who rated at least one item of each factor greater than or equal to four.

Regarding the ESC, 77 questionnaires were returned and 76 were analyzed (for the analysis, a letter from A to T was assigned to each candidate). For the Tue2 session, public reports showed significant differences between candidates for the global factor [$F(5,190) = 5.13$; $p < 0.0001$; $W = 0.12585$] and *post hoc* analysis revealed differences between candidates S ($M = 1.54$; $SD = 1.5$) vs. L ($M = 2.64$; $SD = 1.8$; $p < 0.001$), I ($M = 1.41$; $SD = 1.43$) vs. L ($p < 0.0001$) and L vs. P ($M = 1.26$; $SD = 1.33$; $p < 0.0001$). Candidate L had a higher global factor rating than the other candidates. For the Tue2 session, there were also significant differences for the absorption factor [$F(5,190) = 3.65$; $p < 0.0125$; $W = 0.07608$]. The *post hoc* analysis revealed differences between candidates S ($M = 1.13$; $SD = 1.06$) vs. L ($M = 2$; $SD = 1.52$; $p < 0.01$), I ($M = 1.33$; $SD = 1.36$) vs. L ($p < 0.001$) and L vs. P ($M = 1.08$; $SD = 1.20$; $p < 0.001$). For more details, see **Supplementary Appendix 4**. Candidate L had a higher global factor rating than the other candidates. There were no significant differences for chill factor between candidates [$F(5,190) = 2.04$; $p = 0.069$; $W = 0.06533$] or for the touched factor [$F(5,190) = 2.25$; $p = 0.0464$; $W = 0.07506$]. For Wed1, there were no significant differences between candidates for any factor of the AES-M [touched factor $F(5,180) = 5,180$; $p = 0.696$; $W = 0.01679$; absorption factor $F(5,180) = 1.51$; $p = 0.182$; $W = 0.04666$; chill factor $F(5,180) = 0.70$; $p = 0.626$; $W = 0.0288$] or global factor [$F(5,180) = 1.41$; $p = 0.222$; $W = 0.0486$]. **Table 4** presents the number of people reporting chills (at least one item of the chill factor) by candidate for both Tue2 and Wed1. The phrase "I felt chills for other candidates" indicates the number of people who also indicated experiencing chills for other candidates.

During the Tue2 session, 15 participants (38.4%) indicated feeling chills, and during the Wed1 session, 14 participants (37.8%) indicated feeling chills. During the whole competition (Mon1, Mon2, Tue1, Tue2, and Wed1), 58 participants (34.9%) indicated feeling chills.

Regarding the AFO classification, 32 members of the public completed the self-reported AFO categories. The results indicate that there are significantly more "curious music lovers" compared to the AFO public survey (53.1% vs. 15.7%; $p < 0.0001$; $V = 0.00542$); there are significantly fewer "sociable" respondents in our study than in the AFO survey (9.4% vs. 29.6%; $p < 0.05$; $V = -0.0234$) and there are significantly fewer "distanced amateurs" in our study than in the AFO survey (3% vs. 20.9%; $p < 0.05$; $V = -0.0231$). There are no significant differences between the "occasional layman" and "classical music lover" profiles.

Suitability of Competition Settings

For Mon1 and Mon2, the candidates' performances were timed; they had precisely 10 min. During each performance,

TABLE 4 | The number of people reporting feeling chills, by candidate (Tue2 and Wed1).

Tue2 N = 39						
Candidate	S	E	O	I	L	P
Felt chills, <i>n</i> (%)	7 (17.9)	9 (23.1)	5 (12.8)	3 (7.7)	9 (23.1)	2 (5.1)
Felt chills for other candidates, <i>n</i>	6	6	5	3	5	2
Wed1 N = 37						
Candidate	O	C	R	H	I	F
Felt chills, <i>n</i> (%)	5 (13.5)	7 (18.9)	9 (24.3)	5 (13.5)	5 (13.5)	4 (10.8)
Felt chills for other candidates, <i>n</i>	3	6	4	3	2	3

TABLE 5 | Movements, rehearsal and timing for each candidate's performance during the first semi-final round (Wed1) for the second movement.

Order (candidates)	Duration of the whole performance (mins)	Conducting time for the 2nd movement (mins)	Rehearsal work of the 2nd movement (mins)
First (O)	29	18	From measure 27 (6)
Second (C)	23	15	From measure 27 (6)
Third (F)	23	12	From measure 27 (4)
Fourth (H)	23	15	From measure 42 (5)
Fifth (I)	24	13	From measure 37 (5)
Sixth (F)	17	8	Discussion with the pianist and rehearsal (4)

they conducted two extracts for every session, with at least one common extract. For the first semi-final round (Wed1), the candidates had, on average, 23.8 min (SD = 2.7) to perform. They all conducted the second movement during this round. **Table 5** indicates that, except for candidate F, they all had approximately 5 min for free rehearsals, and all rehearsed the second movement with the orchestra. During all the elimination rounds, a real time video of a front view of each conductor's performances was projected onto a large screen. In this way, the audience and the jury could follow the gestures, skills and facial expressions of the candidates.

The segmented part of the competition should enable suitable monitoring of physiological recordings and an analysis of little portions of signal. We identified some suitable spaces to prepare participants, for the positioning of the computers, and for the monitoring of the recordings. For the theater, recordings on musicians would be more complex because of the reduced space in the pit, but was still feasible.

DISCUSSION

The results of this study show the suitability of using a conductor competition as a testbed to study emotional synchronization within and between groups (music performers and listeners) in a natural setting. The conductor competition provides a natural experimental design: during a session, the same orchestra plays the same extract many times in the same place in front of the same public. Thus, every conductor conveyed a different musical interpretation and different emotional content in a randomized running order that constituted natural variation.

Would People Agree to Wear the Sensors During the Competition?

Interestingly, 91% of the interviewees in the different groups agreed to participate by wearing mobile devices. All the people from the jury and the musicians that were interviewed agreed. Only two people from the public were inconsistent with our approach. This result suggests that our research is well accepted by the public, jury and musicians. Even though the organizers of the competition gave their consent and more than 71% of the interviewed conductors agreed, there were four conductors that refused to wear sensors during the competition. For conductors, it might be complicated to implement neurophysiological recordings in terms of the management of technical constraints and not disturb their performance. They may have been afraid of the devices. Nevertheless, all conductors agreed to participate in a study if the research protocol was acceptable for all candidates and was included in the competition rules. However, these results should be carefully considered: agreements in principle are more easily given when people are not directly engaged in the experiment thereafter.

Did the Public Experience Emotions During the Competition?

More than 95% of the participants indicated feeling touched during the competition, according to AES-M questionnaires. Our analysis of the general emotions reveals that a sufficient proportion of the public (29.2%) were able to feel emotions during the first few sessions of the competition. These results also confirm that in the particular context of the competition, even if the musical pieces are repeated and performed with completely

different interpretations than the public's expectations, people can feel emotions. Furthermore, a high percentage of respondents (34.9%) reported feeling chills during the competition. This result strengthens the validity of our approach to the investigation of strong musical emotions. Our evaluation of the public's emotional sensitivity was positive for the first sessions, so we used the ESC on Tue2 and Wed1 to make deeper comparisons between each candidate and each performance. We used the ESC especially when the number of candidates was reduced to six and thus comparisons were easier to make. We confirmed that some candidates conveyed greater emotions to the public than other candidates. In a future natural experiment, conductors conveying greater emotions should also elicit greater emotional synchronization between the musicians and the audience. Less emotional performances could be used as control. During a single performance, some candidates succeeded in triggering chills in almost a quarter of our respondent sample. This suggests that future experiments in natural conditions might highlight an emotional synchronization between participants during a conductor's performance. In sum, our public was sensitive to musical emotion in the natural context of the competition. Furthermore, people that are not sensitive to musical emotions should be a necessary control group for the measurement of emotional synchronization in the future natural experiment.

Music chills are a well-established marker of musical emotional peaks and a quite large majority of humans should be sensitive to music chills – as high as 86%, according to Panksepp (1995) – while some (anhedonics) are not. Music chills do not occur systematically, and many factors are implicated (personality profile, conditions of listening, musical pieces, social atmosphere, temperature of the room, and so on). Our study shows that music chills occurred in a third of the listeners that completed the self-reports under the specific conditions of the contest. The importance of this rate is certainly debatable, but the sample of respondents was not selected on their predisposition to feel chills. This criterion should probably be introduced in a subsequent research study. The current observation simply shows that chills occurred quite frequently despite the conditions of the conductor competition, with short musical excerpts and frequent interruptions of the orchestra. Considering the future naturalistic experiment, this result indicates that an adequate sample can be selected from the public at the competition.

Could Competition Settings Facilitate Experiments?

Considering the competition setting, our observations of the concert halls' characteristics show that neurophysiological recording can be set up in these natural conditions. During all the sessions, each candidate conducted at least one common extract over a sufficient time period and had at least 20 min of rehearsal (except for one candidate). It would allow statistical comparisons of future neurophysiological data to be made. Thus, neurophysiological recordings could be set up to give enough time to obtain comparable data. As they are not directly

engaged in the competition as jury members or musicians, we can hypothesize that members of the public would be cognitively unfettered and thus more able to concentrate on their emotional state for the continuous reporting. With their stationary position allowing for easy control of technical constraints, the public appear to be most able to endure neurophysiological recording. Our population was homogeneous between the sessions: in the different sessions, no significant differences were highlighted for the rating of each item of the TIPI. Since personality profiles are similar between each session, neurophysiological recordings could be implemented in all phases of the competition. This is a strong argument for the feasibility of a future study.

The natural environment lets the public have the flow of natural auditory and visual information matching with the context (Williamon et al., 2014). Moreover, the issue of a conductor or a music competition, impossible to reproduce in simulated contexts, reinforces the immersive experiment and keeps motivation and attention throughout the sessions. This context allows the public to judge every performance. Considering the musical material, the stimuli are ecological and not simulated. To broadcast repeated stimuli many times for neurophysiological recordings in a laboratory experiment is inappropriate. Furthermore, in the condition of a classical concert, classical orchestras' rehearsals are impossible to reproduce, while numerous repetitions are necessary to make comparisons and thus obtain neurophysiological data with a valid statistical analysis. A conductor or a music competition overcomes this difficulty. A stationary instrumental ensemble, in particular a classical one, performing the same musical score several times in the same conditions limits the uncertainty and limits the influence of different natural variables on the neurophysiological data collection. This specific design sets a fixed frame with natural control of musical stimuli in natural conditions. This is essential in studies using music and EEG, since the different features of music (pitch, tone, beat, rhythm) can influence brainwaves in a different way. The conductors will influence the whole orchestra in different ways, conveying different emotional interpretations of musical pieces and producing the aesthetic quality of the ensemble (Boerner and von Streit, 2007). Finally, the participants' responses are a reliable indicator of emotion that needs to be recorded. The natural feedback of participants should not be influenced by any element of the experiment other than the performance itself.

Limitations

Some biases might be implied. Some members of the public might not have fully understood all the terms of the TIPI scale. This kind of questionnaire is most often used in a general context, so our case study context was not totally appropriate. Moreover, though our response rate was reasonable, the sample size remains too small to draw comprehensive conclusions. Our population is slightly different from the general public for classical music performance: in the context of the competition, our sample has significantly more "curious music lovers" and fewer "sociable" respondents than the AFO public survey. These

characteristics limit the external validity of our results. However the “curious music lovers,” defined as particularly interested in the music offered in the context and searching for new musical works, experiences and emotions, are of particular interest in relation to our emotional approach. Furthermore, as mentioned, some people should not have the same opinion about our approach if that implicate to be directly engaged in the competition instead of giving a simple agreement in principle. Because of the competition timing, we could not gather the opinions of all respondents regarding the wearable devices: this meant we were unable to correlate the agreement of the participants and scales. Finally, since one aim of the study was to evaluate the feasibility of neurophysiological recordings in natural conditions, we did not record audio and video motion that would be appropriate for study of the orchestra and public (Swarbrick et al., 2019). Audio and video motion, combined with neurophysiological recordings, could offer a unique approach in the understanding of emotional synchronization and should be included in future research.

Perspectives

These results are part of a larger work investigating emotional synchronization in groups in a natural framework. Our project gathers many facets of non-verbal human interactions (emotional, physical, aesthetic and sensorimotor) and several kinds of interaction in terms of directional categorization: observation condition, turn-based interaction and continuous interaction (Acquadro et al., 2016) in a particular context. Including both the audience and the musicians in a multidirectional interaction increases the ecological validity; unfortunately, though, it also increases the uncertainty of the ongoing measures (D’Ausilio et al., 2015). However, the recording of EEG hyperscanning and physiological data synchronized with the direct behavioral response of multiple participants on a smartphone through different groups will be a huge source of inter-related data, giving complex and precious information about human interactions in passive and active musical contexts. Some studies using EEG hyperscanning have investigated the synchronization between musicians playing in groups (Babiloni et al., 2006; Lindenberger et al., 2009; Müller et al., 2012, 2013). The technological progress provided by recently developed portable EEG equipment combined with increasingly powerful artifact detection methods (Stone et al., 2018) facilitates realistic experiments, especially those studying several musicians simultaneously (Babiloni et al., 2011, 2012; Müller et al., 2012, 2013, 2018). As previously emphasized, classical musicians performing a musical score in a relatively stationary position allows researchers to record EEG data with a high signal-to-noise ratio (Babiloni et al., 2011; D’Ausilio et al., 2015), as well as a good physiological recording. However, though a few studies have recorded EEG and physiological data during musical performances (Babiloni et al., 2011; Ichihashi et al., 2012; Müller et al., 2012, 2013, 2018), one of the main experimental limits is still the control of motor execution, greatly disrupting the signal. Musicians are often instructed to limit their musical execution movements as much as possible. Considering this difficulty,

musicians’ neurophysiological recordings have to be tested and managed before the competition to ensure good signal and minimal noise.

In a future natural experiment, we aim to recruit regulars and people interested primarily in the competition and not in the scientific study. This experiment in natural conditions may highlight as yet unknown cerebral and physiological mechanisms between an active group of musicians and a passive public. We suggest that good synchronization between musicians might have an influence on public emotions. The higher the emotional synchronization in the musician’s group, the more public members should be synchronized. Such a hypothesis will be assessed by objective and subjective measures.

CONCLUSION

These results confirm that this specific conductor competition is a suitable model to investigate collective emotions in natural conditions using neurophysiological devices. Almost all of the people interviewed agreed in principle to wear a sensor; the competition’s emotional content is appropriate for the study of collective emotions; and the settings of the competition fit our neurophysiological approach. The competition design is similar to a natural standard experimental design. Thus, the conductors conveying different emotional performances constitute the intervention, with a randomized running order. Particular attention must be paid to participants’ musical emotional sensitivity. This competition context – particularly an audience gathering in a concert hall – may highlight a number of undiscovered mechanisms involved in emotional synchronization.

DATA AVAILABILITY STATEMENT

The datasets generated for this study are available on request to the corresponding author.

ETHICS STATEMENT

This research work has been approved as conforming to the French law by an ethics committee. It was classified as a psychology observational study outside of the Jardé law (Article R1121-1 of the French Law Code of Public Health amended by decree n°2017-884 of May 2017). Each participant was informed of the observational nature of the study and gave their explicit consent by answering the investigator’s questions. The questionnaires were anonymous and do not allow the identification of any participants.

AUTHOR CONTRIBUTIONS

LP: conceptualization. TC, AC, DG, CJ, and LP: data curation. TC and GT: formal analysis. TC and LP: writing – original draft. TC, AC, DG, LP, and GT: writing – review and editing.

FUNDING

This study was carried out using our laboratory's own funds. This research received no specific grant from any funding agency in the public, commercial or not-for-profit sectors.

ACKNOWLEDGMENTS

We thank all the volunteers for participating in this study. We are also grateful to Jean Michel Mathe, the

director of the Besançon Franche-Comté International Competition for Young Conductors, and Adèle Bernengo, the director of communications, for their support in this research.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: <https://www.frontiersin.org/articles/10.3389/fpsyg.2019.02954/full#supplementary-material>

REFERENCES

- Acquadro, M. A. S., Congedo, M., and De Ridder, D. (2016). Music performance as an experimental approach to hyperscanning studies. *Front. Hum. Neurosci.* 10:185. doi: 10.3389/fnhum.2016.00242
- Association Française des Orchestres (2014). *Les Publics de l'Orchestre - Quand le Public en Cache un Autre*. Paris: Association Française des Orchestres.
- Babiloni, C., Buffo, P., Vecchio, F., Marzano, N., Del Percio, C., Spada, D., et al. (2012). Brains 'in concert': frontal oscillatory alpha rhythms and empathy in professional musicians. *Neuroimage* 60, 105–116. doi: 10.1016/j.neuroimage.2011.12.008
- Babiloni, C., Vecchio, F., Infarinato, F., Buffo, P., Marzano, N., Spada, D., et al. (2011). Simultaneous recording of electroencephalographic data in musicians playing in ensemble. *Cortex* 47, 1082–1090. doi: 10.1016/j.cortex.2011.05.006
- Babiloni, F., Cincotti, F., Mattia, D., Mattiocco, M., Fallani, F. D. V., Tocci, A., et al. (2006). "Hypermethods for EEG hyperscanning," in *Proceedings of the 2006 International Conference of the IEEE Engineering in Medicine and Biology Society*, New York, NY, 3666–3669.
- Balconi, M., and Vanutelli, M. E. (2017). Cooperation and competition with hyperscanning methods: review and future application to emotion domain. *Front. Comput. Neurosci.* 11:86. doi: 10.3389/fncom.2017.00086
- Blood, A. J., and Zatorre, R. J. (2001). Intensely pleasurable responses to music correlate with activity in brain regions implicated in reward and emotion. *Proc. Natl. Acad. Sci. U.S.A.* 98, 11818–11823. doi: 10.1073/pnas.191355898
- Boerner, S., and von Streit, C. F. (2007). Promoting orchestral performance: the interplay between musicians' mood and a conductor's leadership style. *Psychol. Music* 35, 132–143. doi: 10.1177/0305735607068891
- Burland, K., and Pitts, S. (2014). *Coughing and Clapping: Investigating Audience Experience*, 1st Edn. Abingdon: Routledge.
- Chanda, M. L., and Levitin, D. J. (2013). The neurochemistry of music. *Trends Cogn. Sci.* 17, 179–193. doi: 10.1016/j.tics.2013.02.007
- Colver, M. C., and El-Alayli, A. (2016). Getting aesthetic chills from music: the connection between openness to experience and frisson. *Psychol. Music* 44, 413–427. doi: 10.1177/0305735615572358
- Craig, D. G. (2005). An exploratory study of physiological changes during 'chills' induced by music. *Musicae Sci.* 9, 273–287. doi: 10.1177/102986490500900207
- Craig, P., Cooper, C., Gunnell, D., Haw, S., Lawson, K., Macintyre, S., et al. (2012). Using natural experiments to evaluate population health interventions: new MRC guidance. *J. Epidemiol. Community Health* 66, 1182–1186. doi: 10.1136/jech-2011-200375
- D'Ausilio, A., Novembre, G., Fadiga, L., and Keller, P. E. (2015). What can music tell us about social interaction? *Trends Cogn. Sci.* 19, 111–114. doi: 10.1016/j.tics.2015.01.005
- Desmet, F., Leman, M., Lesaffre, M., and De Bruyn, L. (2010). "Statistical analysis of human body movement and group interactions in response to music," in *Advances in Data Analysis, Data Handling and Business Intelligence. Studies in Classification, Data Analysis, and Knowledge Organization*, eds A. Fink, B. Lausen, W. Seidel, and A. Ultsch, (Berlin: Springer), 399–408. doi: 10.1007/978-3-642-01044-6_36
- Dumas, G., Nadel, J., Soussignan, R., Martinier, J., and Garnero, L. (2010). Inter-brain synchronization during social interaction. *PLoS One* 5:e12166. doi: 10.1371/journal.pone.0012166
- Egermann, H., Pearce, M. T., Wiggins, G. A., and McAdams, S. (2013). Probabilistic models of expectation violation predict psychophysiological emotional responses to live concert music. *Cogn. Affect. Behav. Neurosci.* 13, 533–553. doi: 10.3758/s13415-013-0161-y
- Egermann, H., Sutherland, M. E., Grewe, O., Nagel, F., Kopiez, R., and Altenmüller, E. (2011). Does music listening in a social context alter experience? A physiological and psychological perspective on emotion. *Musicae Sci.* 15, 307–323. doi: 10.1177/1029864911399497
- Goldstein, A. (1980). Thrills in response to music and other stimuli. *Psychobiology* 8, 126–129. doi: 10.3758/bf03326460
- Gosling, S. D., Rentfrow, P. J., and Swann, W. B. (2003). A very brief measure of the Big-Five personality domains. *J. Res. Pers.* 37, 504–528. doi: 10.1016/s0092-6566(03)00046-1
- Grewe, O., Katur, B., Kopiez, R., and Altenmüller, E. (2011). Chills in different sensory domains: frisson elicited by acoustical, visual, tactile and gustatory stimuli. *Psychol. Music* 39, 220–239. doi: 10.1177/0305735610362950
- Grewe, O., Kopiez, R., and Altenmüller, E. (2009a). Chills as an indicator of individual emotional peaks. *Ann. N. Y. Acad. Sci.* 1169, 351–354. doi: 10.1111/j.1749-6632.2009.04783.x
- Grewe, O., Kopiez, R., and Altenmüller, E. (2009b). The chill parameter: goose bumps and shivers as promising measures in emotion research. *Music Percept.* 27, 61–74. doi: 10.1525/mp.2009.27.1.61
- Grewe, O., Nagel, F., Kopiez, R., and Altenmüller, E. (2007). Listening to music as a re-creative process: physiological, psychological, and psychoacoustical correlates of chills and strong emotions. *Music Percept.* 24, 297–314. doi: 10.1525/mp.2007.24.3.297
- Ichihashi, K., Bando, H., and Amakawa, T. (2012). Electrodermal activity of professional pianists: sympathetic arousal in piano performance-. *J. Hum. Sci.* 9, 14–29.
- Juslin, P. N., Liljeström, S., Västfjäll, D., Barradas, G., and Silva, A. (2008). An experience sampling study of emotional reactions to music: listener, music, and situation. *Emotion* 8, 668–683. doi: 10.1037/a0013505
- Koehler, K., and Broughton, M. C. (2017). The effect of social feedback and social context on subjective affective responses to music. *Musicae Sci.* 21, 479–498. doi: 10.1177/1029864916670700
- Lamont, A. (2009). "Strong experiences of music in university students," in *Proceedings of the 7th Triennial Conference of European Society for the Cognitive Sciences of Music (ESCOM 2009)*, Jyväskylä.
- Lamont, A. (2011). University students' strong experiences of music: pleasure, engagement, and meaning. *Musicae Sci.* 15, 229–249. doi: 10.1177/102986491101500206
- Leante, L. (2016). "Observing musicians/audience interaction in North Indian classical music performance," in *Musicians and Their Audiences: Performance, Speech, and Mediation*, eds I. Tsioulakis, and E. Hytönen-Ng, (New York, NY: Routledge), 50–65.
- Liljeström, S., Juslin, P. N., and Västfjäll, D. (2013). Experimental evidence of the roles of music choice, social context, and listener personality in emotional reactions to music. *Psychol. Music* 41, 579–599. doi: 10.1177/0305735612440615
- Lindenberger, U., Li, S.-C., Gruber, W., and Müller, V. (2009). Brains swinging in concert: cortical phase synchronization while playing guitar. *BMC Neurosci.* 10:22. doi: 10.1186/1471-2202-10-22

- McCormick, L. (2009). Higher, faster, louder: representations of the international music competition. *Cult. Sociol.* 3, 5–30. doi: 10.1177/1749975508100669
- McCrae, R. R. (2007). Aesthetic chills as a universal marker of openness to experience. *Motiv. Emot.* 31, 5–11. doi: 10.1007/s11031-007-9053-1
- Meyer, B. D. (1995). Natural and quasi-experiments in economics. *J. Bus. Econ. Stat.* 13, 151–161. doi: 10.1080/07350015.1995.10524589
- Molnar-Szakacs, L., Assuied, V. G., and Overy, K. (2011). *Shared Affective Motion Experience (SAME) and Creative, Interactive Music Therapy*. Retrieved from *Musical Imaginations: Multidisciplinary Perspectives on Creativity, Performance and Perception*. Oxford: Oxford University Press.
- Mori, K., and Iwanaga, M. (2014). Resting physiological arousal is associated with the experience of music-induced chills. *Int. J. Psychophysiol.* 93, 220–226. doi: 10.1016/j.ijpsycho.2014.05.001
- Müller, V., Lindenberger, U., and Sängler, J. (2012). Intra- and interbrain synchronization and network properties when playing guitar in duets. *Front. Hum. Neurosci.* 6:312. doi: 10.3389/fnhum.2012.00312
- Müller, V., Sängler, J., and Lindenberger, U. (2013). Intra- and inter-brain synchronization during musical improvisation on the guitar. *PLoS One* 8:e73852. doi: 10.1371/journal.pone.0073852
- Müller, V., Sängler, J., and Lindenberger, U. (2018). Hyperbrain network properties of guitarists playing in quartet. *Ann. N. Y. Acad. Sci.* 1423, 198–210. doi: 10.1111/nyas.13656
- Nieminen, S., Istók, E., Brattico, E., and Tervaniemi, M. (2012). The development of the aesthetic experience of music: Preference, emotions, and beauty. *Musicae Sci.* 16, 372–391. doi: 10.1177/1029864912450454
- Nusbaum, E., Silvia, P., Beaty, R., Burgin, C., Hodges, D., and Kwapil, T. (2013). Listening between the notes: aesthetic chills in everyday music listening. *Psychol. Aesthet. Creat. Arts* 8, 104–109. doi: 10.1037/a0034867
- Nusbaum, E. C., and Silvia, P. J. (2011a). Are openness and intellect distinct aspects of openness to experience? A test of the O/I model. *Pers. Individ. Differ.* 51, 571–574. doi: 10.1016/j.paid.2011.05.013
- Nusbaum, E. C., and Silvia, P. J. (2011b). Shivers and timbres: personality and the experience of chills from music. *Soc. Psychol. Pers. Sci.* 2, 199–204. doi: 10.1177/1948550610386810
- Overy, K. (2012). Making music in a group: synchronization and shared experience. *Ann. N. Y. Acad. Sci.* 1252, 65–68. doi: 10.1111/j.1749-6632.2012.06530.x
- Panksepp, J. (1995). The emotional sources of 'chills'. Induced by music. *Music Percept.* 13, 171–207. doi: 10.2307/40285693
- Sachs, M. E., Ellis, R. J., Schlaug, G., and Loui, P. (2016). Brain connectivity reflects human aesthetic responses to music. *Soc. Cogn. Affect. Neurosci.* 11, 884–891. doi: 10.1093/scan/nsw009
- Salimpoor, V. N., Benovoy, M., Larcher, K., Dagher, A., and Zatorre, R. J. (2011). Anatomically distinct dopamine release during anticipation and experience of peak emotion to music. *Nat. Neurosci.* 14, 257–262. doi: 10.1038/nn.2726
- Salimpoor, V. N., Benovoy, M., Longo, G., Cooperstock, J. R., and Zatorre, R. J. (2009). The rewarding aspects of music listening are related to degree of emotional arousal. *PLoS One* 4:e7487. doi: 10.1371/journal.pone.0007487
- Salkind, N. J. (2010). *Encyclopedia of Research Design*. Thousand Oaks, CA: SAGE Publications Inc.
- Schoeller, F. (2015). Knowledge, curiosity, and aesthetic chills. *Front. Psychol.* 6:1546. doi: 10.3389/fpsyg.2015.01546
- Shoda, H., Adachi, M., and Umeda, T. (2016). How live performance moves the human heart. *PLoS One* 11:e0154322. doi: 10.1371/journal.pone.0154322
- Shoval, N., Schwimer, Y., and Tamir, M. (2018). Real-time measurement of tourists' objective and subjective emotions in time and space. *J. Travel Res.* 57, 3–16. doi: 10.1177/0047287517691155
- Silvia, P., and Nusbaum, E. (2011). On personality and piloerection: individual differences in aesthetic chills and other unusual aesthetic experiences. *Psychol. Aesthet. Creat. Arts* 5, 208–214. doi: 10.1037/a0021914
- Stone, D. B., Tamburro, G., Fiedler, P., Haueisen, J., and Comani, S. (2018). Automatic removal of physiological artifacts in EEG: the optimized fingerprint method for sports science applications. *Front. Hum. Neurosci.* 12:96. doi: 10.3389/fnhum.2018.00096
- Sumpf, M., Jentschke, S., and Koelsch, S. (2015). Effects of aesthetic chills on a cardiac signature of emotionality. *PLoS One* 10:e0130117. doi: 10.1371/journal.pone.0130117
- Swarbrick, D., Bosnyak, D., Livingstone, S. R., Bansal, J., Marsh-Rollo, S., Woolhouse, M. H., et al. (2019). How live music moves us: head movement differences in audiences to live versus recorded music. *Front. Psychol.* 9:2682. doi: 10.3389/fpsyg.2018.02682
- Tarr, B., Launay, J., and Dunbar, R. I. M. (2014). Music and social bonding: 'self-other' merging and neurohormonal mechanisms. *Front. Psychol.* 5:1096. doi: 10.3389/fpsyg.2014.01096
- Toppi, J., Borghini, G., Petti, M., He, E. J., Giusti, V. D., He, B., et al. (2016). Investigating cooperative behavior in ecological settings: an EEG hyperscanning study. *PLoS One* 11:e0154236. doi: 10.1371/journal.pone.0154236
- Volpe, G., D'Ausilio, A., Badino, L., Camurri, A., and Fadiga, L. (2016). Measuring social interaction in music ensembles. *Philos. Trans. R. Soc. Lond. B. Biol. Sci.* 371:20150377. doi: 10.1098/rstb.2015.0377
- Wilhelm, F. H., and Grossman, P. (2010). Emotions beyond the laboratory: theoretical fundamentals, study design, and analytic strategies for advanced ambulatory assessment. *Biol. Psychol.* 84, 552–569. doi: 10.1016/j.biopsycho.2010.01.017
- Williamon, A., Aufegger, L., and Eiholzer, H. (2014). Simulating and stimulating performance: introducing distributed simulation to enhance musical learning and performance. *Front. Psychol.* 5:25. doi: 10.3389/fpsyg.2014.00025

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

Copyright © 2020 Chabin, Tio, Comte, Joucla, Gabriel and Pazart. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.



Cortical Patterns of Pleasurable Musical Chills Revealed by High-Density EEG

Thibault Chabin^{1*}, Damien Gabriel^{1,2,3}, Tanawat Chansophonkul², Lisa Michelant¹, Coralie Joucla¹, Emmanuel Haffen^{1,2,3}, Thierry Moulin^{1,2,3}, Alexandre Comte^{1,2,3} and Lionel Pazar^{1,2,3}

¹ Laboratoire de Neurosciences Intégratives et Cliniques, EA 481, Université Bourgogne Franche-Comté, Besançon, France, ² INSERM CIC 1431, Centre d'Investigation Clinique de Besançon, Centre Hospitalier Universitaire de Besançon, Besançon, France, ³ Plateforme de Neuroimagerie fonctionnelle et Neurostimulation – Neuraxess, Centre Hospitalier Universitaire de Besançon, Besançon, France

Music has the capacity to elicit strong positive feelings in humans by activating the brain's reward system. Because group emotional dynamics is a central concern of social neurosciences, the study of emotion in natural/ecological conditions is gaining interest. This study aimed to show that high-density EEG (HD-EEG) is able to reveal patterns of cerebral activities previously identified by fMRI or PET scans when the subject experiences pleasurable musical chills. We used HD-EEG to record participants (11 female, 7 male) while listening to their favorite pleasurable chill-inducing musical excerpts; they reported their subjective emotional state from low pleasure up to chills. HD-EEG results showed an increase of theta activity in the prefrontal cortex when arousal and emotional ratings increased, which are associated with orbitofrontal cortex activation localized using source localization algorithms. In addition, we identified two specific patterns of chills: a decreased theta activity in the right central region, which could reflect supplementary motor area activation during chills and may be related to rhythmic anticipation processing, and a decreased theta activity in the right temporal region, which may be related to musical appreciation and could reflect the right superior temporal gyrus. The alpha frontal/prefrontal asymmetry did not reflect the felt emotional pleasure, but the increased frontal beta to alpha ratio (measure of arousal) corresponded to increased emotional ratings. These results suggest that EEG may be a reliable method and a promising tool for the investigation of group musical pleasure through musical reward processing.

Keywords: high density EEG, EEG, emotion, music, chills, musical reward, peak pleasure, cerebral activity

INTRODUCTION

The power of music over human emotions is intriguing and there is an ongoing debate regarding not the mechanisms of how music can provoke pleasure but rather why music can be a rewarding experience (Goupil and Aucouturier, 2019). Music causes both marked changes in an individual's emotional state to the point of musical thrill (Goldstein 1980) and collective emotional contagion in a social context (Egermann et al., 2011). At the individual level, intense musical emotion involving areas of the limbic system can provoke a pleasurable rewarding experience: the musical

OPEN ACCESS

Edited by:

Eckart Altenmüller,
Hanover University of Music, Drama
and Media, Germany

Reviewed by:

Pablo Ripolles,
New York University, United States
Dan Zhang,
Tsinghua University, China

*Correspondence:

Thibault Chabin
tchabin@edu.univ-fcomte.fr;
chabinthibault@yahoo.fr

Specialty section:

This article was submitted to
Auditory Cognitive Neuroscience,
a section of the journal
Frontiers in Neuroscience

Received: 26 May 2020

Accepted: 29 September 2020

Published: xx October 2020

Citation:

Chabin T, Gabriel D,
Chansophonkul T, Michelant L,
Joucla C, Haffen E, Moulin T,
Comte A and Pazar L (2020)
Cortical Patterns of Pleasurable
Musical Chills Revealed by
High-Density EEG.
Front. Neurosci. 14:565815.
doi: 10.3389/fnins.2020.565815

chill (Blood and Zatorre, 2001; Salimpoor et al., 2009, 2011). In their paradigm, the participants provided highly pleasurable musical extracts and continuously reported their felt emotional pleasure while listening to music (neutral, low pleasure, high pleasure, and chills). The phenomenon of pleasure associated with chills is composed of two phases, an anticipation phase before the chill when the pleasure is growing, with a specific dopamine release in the dorsal striatum (caudate), and a peak pleasure phase with a dopaminergic release in the ventral striatum (nucleus accumbens) (Salimpoor et al., 2011). In addition, a pharmacological study by Ferreri et al. (2019) demonstrated that dopaminergic releases were not only the consequence but actually one of the causes of the felt emotional pleasure. These cerebral investigations remain limited to laboratory experiments with heavy neuroimaging techniques (fMRI, PET scan) whereas social neurosciences, and consequently the study of collective emotions, is moving toward natural/ecological paradigms (Acquadro et al., 2016; Volpe et al., 2016; Dikker et al., 2017; Bevilacqua et al., 2018; Soto et al., 2018; Swarbrick et al., 2018; Matusz et al., 2019). The use of mobile wireless EEG could provide the possibility of studying cerebral activity during peak emotion of musical chills in ecological/natural conditions, especially because it can be used in hyperscanning paradigms with several participants simultaneously (Lindenberger et al., 2009; Müller et al., 2013, 2018; Sanger et al., 2013; Acquadro et al., 2016).

EEG is a direct measure of electrical activity of the brain that confers a high temporal resolution in the millisecond range. With high-density EEG (HD-EEG), a large area of the scalp is covered by electrodes, which makes it possible to investigate the cortical sources of the surface activity with an acceptable spatial resolution (estimated in the range of a centimeter). Although musical chills have not yet been studied by EEG, HD-EEG enables the source reconstruction to identify the cortical origins of surface activities (Michel and Brunet, 2019). The activities of these sources can therefore be compared with activities already known in structures implicated in musical reward.

Earlier data from classic EEGs suggest that low frequencies such as theta oscillations are involved in reward processing (Aftanas and Golocheikine, 2001; Wingerden et al., 2010; Gruber et al., 2013). Theta activity is reported to be linked to memory processing in reward contexts (Gruber et al., 2013; Pu and Yu, 2019). A stronger activity in the theta frequency range was found in fronto-medial sites after rewarding feedback onset (Pornpattananangkul and Nusslock, 2016) and at occipito-parietal and fronto-central electrodes when comparing anticipation of high versus low reward (Steiger and Bunzeck, 2017). In the musical context, theta band changes in the fronto-central area during music appreciation have been reported and also correspond to high emotional arousal (Lin et al., 2010). Compared with neutral musical excerpts, highly pleasant music provokes a higher oscillatory activity in the theta band over fronto-central regions than neutral musical excerpts (Nemati et al., 2019). The theta power increases in the medial frontal area for positive musical stimuli (Omicie et al., 2015) and for pleasant music (Sammler et al., 2007). The more the loops of the anterior cingulate cortex (ACC) and the medial frontal cortex

are activated while listening to music, the more the fronto-medial theta power increases (Jancke et al., 2015). The specific theta/alpha activities relative to the valence of the music (both aversive/negative and attractive/positive) might be mediated by the amygdala, which is a direct modulator of the auditory and orbitofrontal cortex for processing musical emotion (Omicie et al., 2015). It has also been suggested that theta oscillations play a role in the synchronization of temporal and prefrontal structures during musical emotional processing (Lewis, 2005). Furthermore, the overall frontal alpha EEG asymmetry (power difference between right and left hemisphere), which is reported to be a good indicator of the emotional state while listening to music (Schmidt and Trainor, 2001; Davidson, 2004; Harmon-Jones et al., 2010; Ramirez and Vamvakousis, 2012; Vecchiato et al., 2012; Arjmand et al., 2017; Mennella et al., 2017; Ramirez et al., 2018), should be an interesting candidate for the study of cortical patterns involved during musical chills.

The aim of the present study was to use HD-EEG to identify specific cortical patterns that underlie musical chills. Our main hypothesis was that an increased felt pleasure on a scale from neutral (lowest intensity) to chills (maximum intensity) should produce an increase of the theta activity over the fronto-central regions, and a specific pattern of theta activity over the centroparietal and temporal regions.

MATERIALS AND METHODS

Ethics

The study was approved by an independent ethics committee (CPP Ouest V – Rennes; no. 2018-A01653-52) and follows recommendations from the French Jard’s Law (Article R1121-11 of the French Law Code of Public Health amended by decree 127 no. 2017-884 of May 2017) on non-invasive protocols involving healthy humans. All participants received the information in full both orally and on paper, and signed a written informed consent form before inclusion in the study.

Participants

The sample for this study consisted of 18 healthy volunteers (11 women, 7 men) with a mean age of 39.7 years (SD 18.3, range 18–73). All were right-handed (Edinburgh Inventory score >50; Oldfield, 1971), were sensitive to musical reward according to the Barcelona Music Reward Questionnaire (Mas-Herrero et al., 2013; Saliba et al., 2016) (BMRQ overall score >65, Hedonic/High-hyper hedonic group from Mas-Herrero et al., 2014), and frequently experienced chills induced by pleasurable music. Ten participants in this sample were amateur musicians (mean of 20.2 years of practice, SD 13.37, range 10–55).

Recruitment was carried out via advertising posters in our university and hospital. Eighty-nine people responded to our request or contacted us directly to participate in this study. Fifty-eight did not fulfill the eligibility criteria (2 scored lower than 65 for the BRMQ, 3 were not exclusively right-handed, 9 did not provide enough chill-inducing extracts, 36 did not respond after their initial request for information/did not send back questionnaires, for 7 we could not program an

229 appointment, and 1 had a haircut incompatible with EEG
230 recording). Eight participants were not re-contacted after their
231 request to participate because the inclusion deadline had passed.
232 Finally, 23 subjects passed a medical examination to confirm they
233 had normal hearing (with audiogram tests), and an absence of
234 neurological or psychiatric disorders. They all signed an informed
235 consent form. Among them, five participants were ultimately not
236 included in the analysis; three participants reported one or zero
237 instances of chills during the experiment, and for two participants
238 not enough EEG epochs free from artifacts were recorded.

239 EEG Recordings

241 EEG signals were recorded using a 256-channel Sensor Net
242 (from Electrical Geodesics) with a NetAmp 300 high impedance
243 amplifier (Electrical Geodesics). Continuous recordings were
244 performed with a high-pass set at 0.1 Hz and a sampling rate
245 of 1000 Hz; all channels were referenced to the vertex (Cz) and
246 impedances were below 50 k Ω . Cartool software (version 3.7) was
247 used to pre-process all EEG data, a notch filter fixed to 50 Hz was
248 applied, data were band-pass filtered between 1 and 30 Hz, and
249 an average re-referencing of all channels was performed.

250 Procedure and Behavioral Data

252 During the listening session, participants were seated in a
253 comfortable chair, kept their eyes closed, and listened via intra-
254 auricular headphones (EarPods) to five chill-inducing extracts
255 (without any restriction regarding musical style or musical trend)
256 that they had provided and three additional neutral extracts
257 selected by experimenters. They were asked to continuously
258 report their felt emotional pleasure by continuously pushing on
259 a four-button response box according to four levels: (1) neutral,
260 (2) low pleasure, (3) high pleasure, and (4) chills (from Salimpoor
261 et al., 2009, 2011).

262 The extracts were separated by 30-s pauses without music,
263 giving a listening session of approximately 15 min total (see
264 schema of experiment in Figure 1). All musical extracts were
265 cut to 90 s, including 60 s before the peak pleasure (indicated
266 by the participants before the experiment themselves), and were
267 normalized to 0 dB including a progressive fade in/fade out of 3 s.

268 The neutral extracts were selected from a list of musical
269 extracts that had been rated for attractiveness/averseness and
270 arousal by an independent sample of 12 people before the
271 experiment. The mean attractiveness/averseness of the three final
272 selected extracts was between -1 and $+1$ (on a scale from
273 $-5 =$ “aversive extract” to $+5 =$ “very pleasurable”) and the mean
274 arousal was lower than 3 (on a scale from 0 = “low intensity” to
275 10 = “max intensity”). The participants included in the study were
276 not familiar with these neutral extracts.

277 Three listening sessions were recorded using three different
278 EEG systems. Only one recording session used HD-EEG, and the
279 results presented in this paper concern only the HD-EEG data.
280 Participants were asked to report each chill event when they felt
281 high emotional pleasure accompanied by a physical sensation
282 such as goose bumps or thrill, hair standing on end, tingling
283 sensations, or shivers down the spine. If they were already
284 pushing the chills button, they were asked to release and press the
285 chills button again to indicate each “new chill.” Each participant

286 rated their general enjoyment of the experiment (regarding
287 experimental conditions and the experimental procedure) on a
288 10-point Likert scale from 1 (I did not enjoy the experiment at
289 all) to 10 (I really enjoyed the experiment) and rated each of the
290 own excerpts for arousal (from 1; “this extract is very calming”
291 to 10; “this extract is very exciting/arousing”) and emotional
292 valence (from -5 ; “this extract has a negative valence, it evoked
293 sad feelings/melancholia, etc.” to $+5$; “this extract has a positive
294 valence, it evoked joy, happiness etc”).

295 Analysis and Statistics

297 The neutral extracts were chosen to be neither pleasant
298 nor unpleasant. Most of the participants reported difficulty
299 distinguishing between the neutral and low pleasure emotional
300 states and usually used only one of the two conditions when
301 reporting low emotional state. These two conditions were
302 therefore grouped together as the “Low pleasure” condition.
303 We extracted the periods of time for which the participant
304 reported each of low pleasure, high pleasure, or chills. Epochs
305 of 1 s were extracted for these periods for the three emotional
306 conditions: (1) low pleasure, (2) high pleasure, and (3) chills.
307 Every epoch with visually detectable artifacts was removed
308 from the analysis. A frequency analysis using fast Fourier
309 transform (FFT) was performed on three frequency bands for
310 each epoch of each condition, theta (4–8 Hz), alpha (8–12 Hz),
311 and beta (12–20 Hz), using MATLAB software (2019a). The
312 power spectral density (PSD) was estimated for each condition
313 using Welch’s method (Hanning window 50% overlap) based
314 on FFT magnitude squared. Eight regions of interest (ROIs)
315 were investigated for each frequency band: left central (LC),
316 right central (RC), left frontal (LF), right frontal (RF), left pre-
317 frontal (LPF), right pre-frontal (RPF), left temporal (LT), and
318 right temporal (RT) (see Figure 2). A calculation of valence
319 was performed following approach/withdrawal theory (Schmidt
320 and Trainor, 2001; Altenmüller et al., 2002; Arjmand et al.,
321 2017); the mean inter-hemispheric PSD value difference in the
322 alpha frequency range was calculated for the frontal area and
323 pre-frontal area as follows:

$$324 \text{Alpha frontal asymmetry} = (\alpha \text{ PSD RF ROI}) -$$

$$325 (\alpha \text{ PSD LF ROI})$$

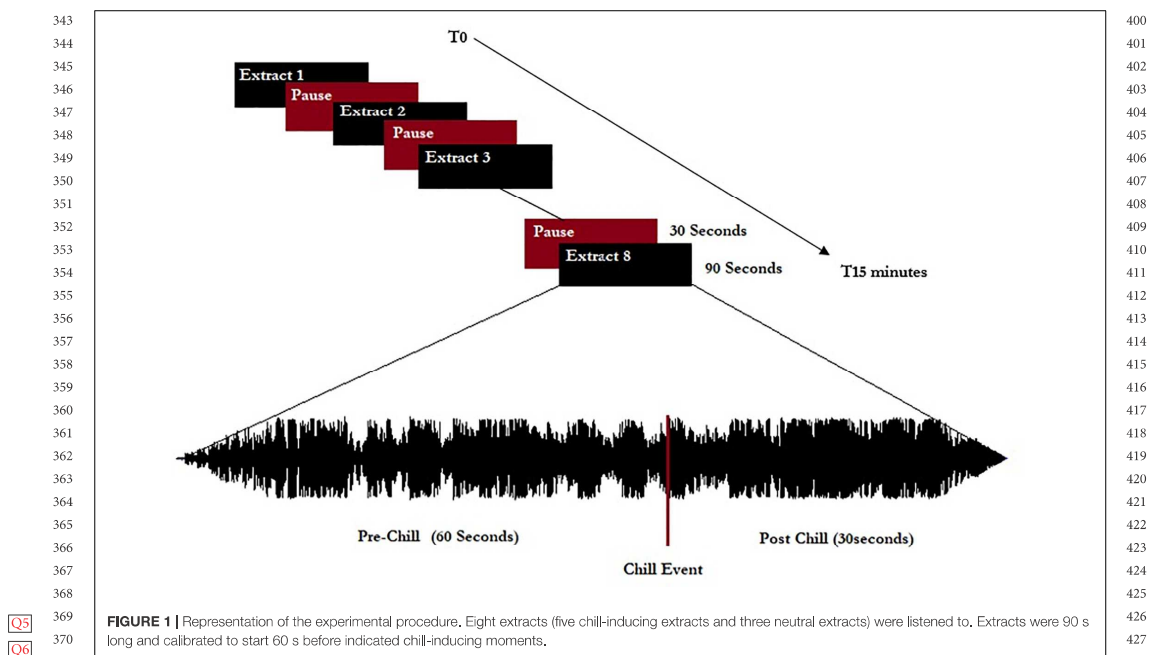
$$326 \text{Alpha prefrontal asymmetry} = (\alpha \text{ PSD RPF ROI}) -$$

$$327 (\alpha \text{ PSD LPF ROI})$$

328 An analysis of arousal based on the method used by Ramirez
329 and Vamvakousis (2012) and Ramirez et al. (2018) was performed
330 by calculating the beta to alpha ratio according to the following
331 equation:

$$332 \text{Beta to alpha ratio} = \frac{(\beta \text{ PSD ROI LF} + \beta \text{ PSD ROI RF})}{333 (\alpha \text{ PSD ROI LF} + \alpha \text{ PSD ROI RF})}$$

334 To identify the source that generates the activities observed
335 on the scalp, it is possible to apply source localization analyses
336 that convert EEGs to a real neuroimaging modality (Michel
337 and Brunet, 2019). Based on the structures and regions that are
338



QS
Q6

already known to be involved in musical pleasure and music chills, the identification of sources of signal would reinforce the interpretation of surface activities. Source localization was performed on the mean participant's grand averages of 1-s epochs for each condition. The source localization was estimated using a linear inverse solution based on a local autoregressive average (LAURA) model (Menendez et al., 2001; Brunet et al., 2011). Source localization was applied to five regions of interest, which were created using Cartool software and were based on fMRI literature (Blood and Zatorre, 2001; Brown et al., 2004; Salimpoor et al., 2011): (1) orbito-frontal gyrus (OFG), (2) supplementary motor area (SMA), (3) bilateral insula (Ins), (4) right superior temporal gyrus (RSTG), and (5) left superior temporal gyrus (LSTG).

All statistical analysis was performed using R Studio software (version 3.5.2; 2018-12-20). For behavioral data, Pearson r correlations were performed to explore the link between the number of chills reported and BMRQ scores, sex, age, and years of musical practice. The normality of each data set was confirmed using Kolmogorov–Smirnov tests. For EEG analysis, three conditions (Low pleasure, High pleasure, and Chills) were compared for *alpha frontal asymmetry*, *alpha prefrontal asymmetry*, *beta to alpha ratio*, and for PSD values in the theta and alpha frequency bands for each ROI (LC, RC, LF, RF, LPE, RPE, LT, and RT) using repeated measures ANOVA ($\alpha = 0.05$) followed by *post hoc* tests using the Bonferroni correction within ROIs. For source localization, repeated measures ANOVA were applied

to compare the three conditions within ROIs (SMA, OFC, Ins, LSTG, and RSTG) using grand averages of epochs, and *post hoc* tests were performed using the Bonferroni correction based on the multiplicity of the tests within and between ROIs.

RESULTS

Behavioral Data

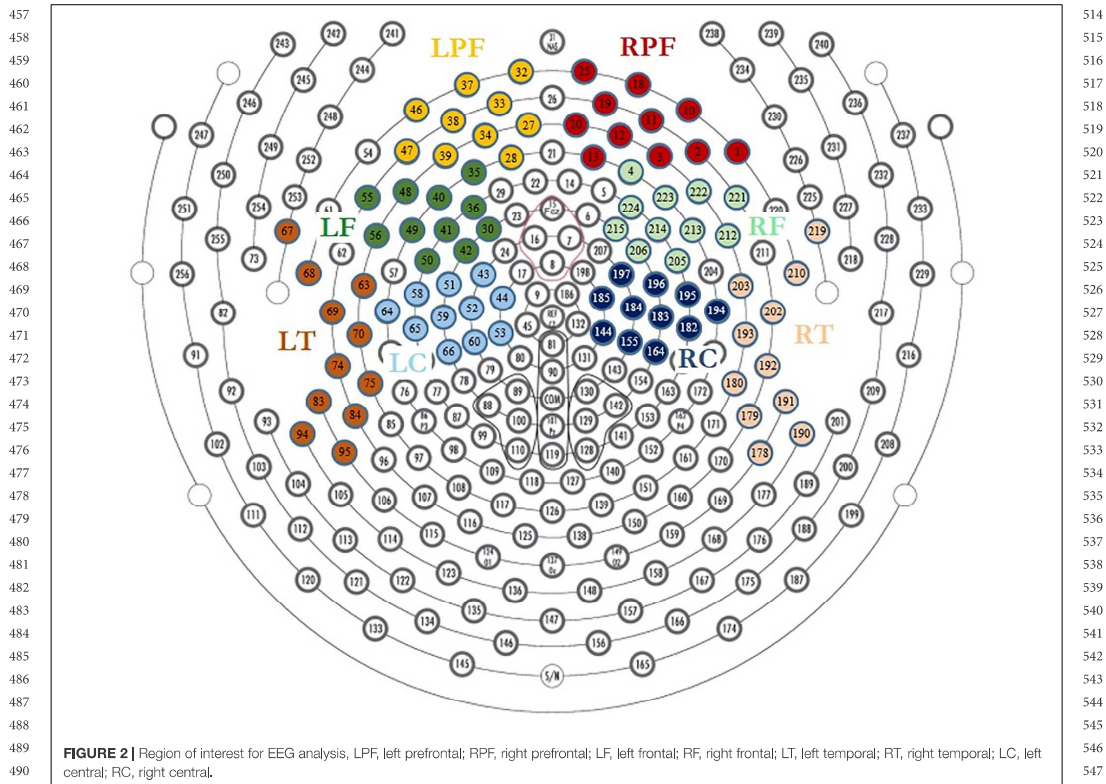
The 18 participants reported 305 chills during the listening sessions, with a mean of 16.9 (SD 12.8) chills per participant and a mean duration of 8.75 s (SD 10.71) (details in **Table 1**). The durations of reported chills were very variable across participants. Many chills occurred outside the times of expected peak pleasure. Our analysis revealed no relationship between the number of chill reports and sex, age, BMRQ score, or number of years of music practice. The participants rated their overall enjoyment of the experiment at a mean of 7.9 (SD 0.9, range 5–9) on the 10-point Likert scale. It appears that chills reported during the listening by participants did not always match with the peak pleasure previously indicated by the participant.

EEG Results

Power Spectral Density Value

Theta frequency band

The results show specific activity in the theta frequency range in the central, right prefrontal, and right temporal regions. The



comparison of PSD values in the theta frequency range for each condition (Low pleasure, High pleasure, and Chills) revealed a significant group effect for the LC ROI [$F(2,15) = 3.75$; $p = 0.033$], RC ROI [$F(2,15) = 4.09$; $p = 0.025$], RT ROI [$F(2,15) = 5.88$; $p = 0.006$], and RPF ROI [$F(2,15) = 3.28$; $p = 0.049$] (Figure 3). *Post hoc* analysis revealed that PSD values were significantly lower for Chills versus Low pleasure (RC ROI $p = 0.042$; RT ROI $p = 0.004$) and the PSD values were significantly higher for Chills versus Low pleasure for the RPF ROI ($p = 0.046$).

Post hoc analysis did not reveal any differences between Chills versus Low pleasure for the LC ROI ($p = 0.058$), or any significant differences between Low and High pleasure for LC ROI ($p > 1$), RC ROI ($p > 1$), RT ROI ($p = 0.3$), and RPF ROI ($p = 0.86$). *Post hoc* analysis also did not reveal significant differences between High pleasure and Chills for the LC ROI ($p = 0.085$), RC ROI ($p = 0.072$), RT ROI ($p = 0.27$), or RPF ROI ($p = 0.44$).

There were no significant effects between the different conditions but rather “tendencies” showing an increase in the theta PSD value for the LPF ROI [$F(2,15) = 2.84$; $p = 0.071$], and a decrease in the theta PSD value for the LF ROI [$F(2,15) = 2.78$;

$p = 0.075$] and RF ROI [$F(2,15) = 3.15$; $p = 0.055$] as the emotional rating increased (Figure 3). There were no effects for the LT ROI [$F(2,15) = 0.13$; $p = 0.87$].

Alpha frequency band

The same analysis comparing the PSD value in alpha frequency bands for each condition did not reveal significant group effects for any ROI (all p -values > 0.9) (Figure 3).

Beta to Alpha Ratio and Alpha Asymmetry

For *beta to alpha ratio* analysis, the comparison of each condition showed a significant group effect [$F(2,15) = 4.77$; $p = 0.014$] (Figure 4). *Post hoc* analysis revealed significant differences between Chills and the two other conditions (Chills vs. Low Pleasure, $p = 0.041$; Chills vs. High Pleasure, $p = 0.028$). The *beta to alpha ratio* was higher for chills than for low pleasure and high pleasure.

For asymmetry analysis, the comparison of each condition did not show a significant group effect either for *alpha frontal asymmetry* [$F(2,15) = 0.32$; $p = 0.75$] or for *alpha prefrontal asymmetry* [$F(2,15) = 0.06$; $p = 0.94$] (Figure 4).

571 **TABLE 1** | Number of reported chills and mean duration by participants.

572 Participant	573 Overall BMRQ score	574 Number of reported chills	575 Number of selected musical extracts with at least 1 chill reported (min-max number of chills by extract)	576 Mean duration of reported chills in seconds (SD)
577 1	69	15	3 (0–8)	1.40 (0.75)
578 2	89	17	5 (1–5)	2.14 (1.01)
579 3	88	19	5 (2–7)	4.22 (6.55)
580 4	88	16	5 (1–4)	1.72 (0.76)
581 5	85	19	5 (1–6)	0.2 (0.27)*
582 6	91	48	5 (5–15)	2.27 (0.78)
583 7	83	8	5 (1–3)	4.25 (2.41)
584 8	76	11	5 (1–2)	11.19 (10.21)
585 9	81	6	5 (1–2)	39.77 (24.29)
586 10	72	7	4 (1–2)	4.11 (4.67)
587 11	79	9	5 (1–2)	31.96 (21.76)
588 12	84	4	4 (1–1)	15.92 (9.54)
589 13	73	14	5 (2–5)	10.06 (9.09)
590 14	91	13	5 (1–4)	5.2 (2.02)
591 15	72	8	2 (1–7)	4.88 (2.19)
592 16	78	50	5 (8–13)	3.45 (1.48)
593 17	84	17	5 (2–6)	5.26 (2.60)
594 18	83	24	5 (1–8)	5.18 (4.49)

595 *Participants five indicated only the beginning of the felt chills and did not maintain the button pressure.

601 Source Imaging

602 Source localization revealed a positive link between increased
603 emotional rating and intensity of activation of the medial orbito-
604 frontal cortex [$F(2,15) = 17.4$; $p < 1.10^{-5}$], the bilateral insula
605 [$F(2,15) = 21.63$; $p < 1.10^{-6}$], the supplementary motor area
606 [$F(2,15) = 27.3$; $p < 1.10^{-7}$], and the left and right superior
607 temporal gyri [LSTG: $F(2,15) = 17.76$, $p < 0.00001$; RSTG:
608 $F(2,15) = 22.05$, $p < 1.10^{-6}$]. Statistical analysis revealed a
609 significant difference for Chills versus Low pleasure (OFC:
610 $p < 0.001$, Ins: $p < 0.0001$, SMA: $p < 1.10^{-5}$, LSTG: $p < 0.001$,
611 RSTG: $p < 0.0001$) and for Chills versus High pleasure (OFC:
612 $p < 0.01$, Ins: $p < 0.001$, SMA: $p < 0.0001$, LSTG: $p < 0.01$, RSTG:
613 $p < 0.001$) (see **Figure 5**).

614 There were no significant differences when comparing all the
615 ROIs for High versus Low pleasure (all p -values > 1).

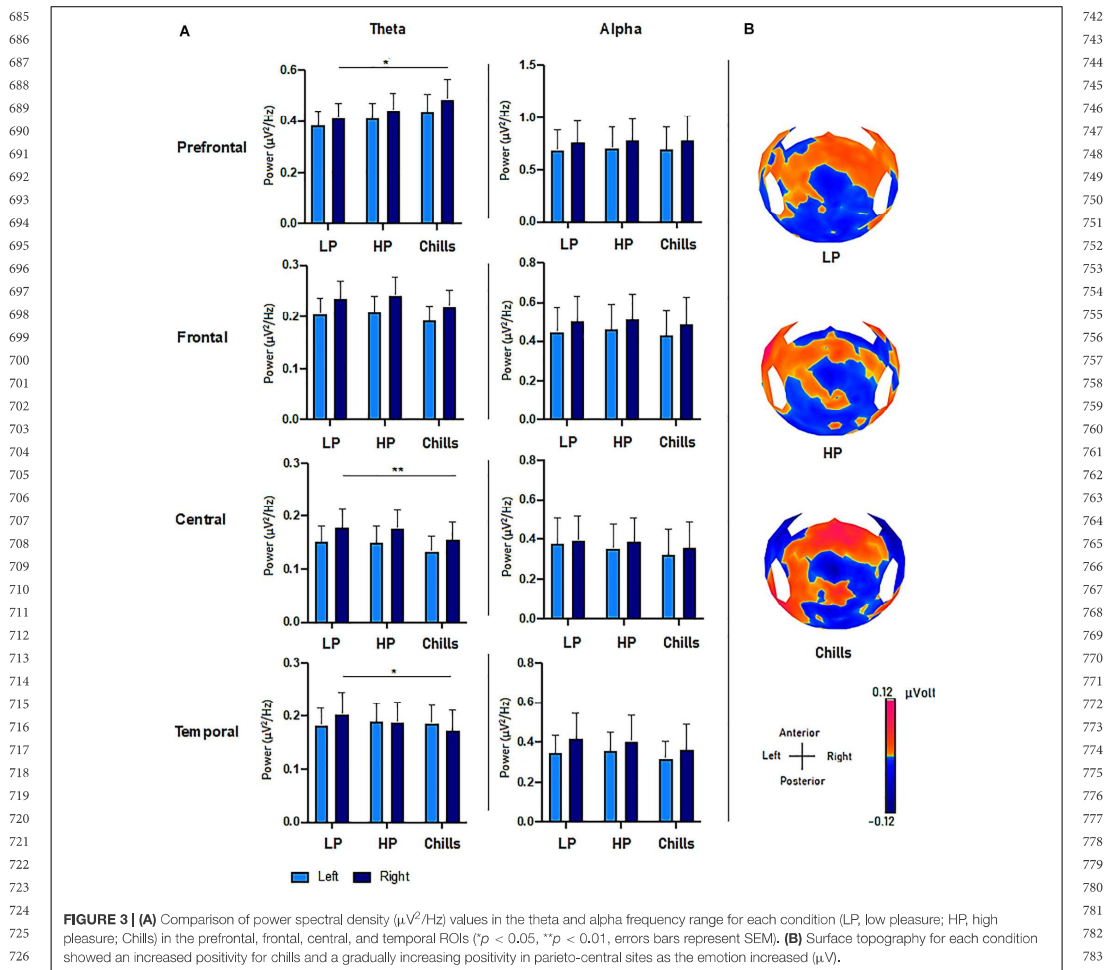
618 GENERAL DISCUSSION

620 The aim of the present study was to investigate the neural
621 processing that occurs during musical chills using HD-EEG. The
622 EEG results confirmed a specific activity over the prefrontal
623 regions in the theta frequency band, but we observed no specific
624 alpha asymmetry patterns relating to the approach/withdrawal
625 system. Source localization allowed us to identify greater activities
626 in the insula, OFC, and SMA, which confirmed the specificity
627 of surface activity for chills. These three activated areas are

628 in line with previous findings of neuroimaging studies that
629 also identified the bilateral insula, OFC, ventromedial prefrontal
630 cortex, SMA, and ventral/dorsal striatum involved in musical
631 chills (Blood et al., 1999; Blood and Zatorre, 2001; and Brown
632 et al., 2004).

633 The specific theta activity has been previously identified in
634 musical emotional processing (Ramos and Corsi-Cabrera, 1989).
635 The results from the prefrontal areas confirm previous findings
636 that suggested a higher theta band activity over the fronto-
637 midline regions and fronto-medial region for pleasant versus
638 neutral musical excerpts (Sammler et al., 2007; Jäncke et al.,
639 2015; Nematı et al., 2019), for positive musical stimuli (Omigie
640 et al., 2015; Rogenmoser et al., 2016). Theta activity might be
641 linked with reward anticipation, memory, and attention (Burgess
642 and Gruzelier, 1997; Krause et al., 1999; Gruber et al., 2013;
643 Pu and Yu, 2019), which is a key component of emotional
644 responses to music (Brattico and Pearce, 2013). More precisely,
645 theta activity is associated with successful memory performance
646 in context of high reward and is reported to be linked with
647 dopaminergic activity (Gruber et al., 2013). The specific frontal
648 theta activity could be the surface activity reflection of the
649 reward system activation via a network including the amygdala,
650 insula, and OFC, which has previously been found to correspond
651 to the occurrence of chills (Blood and Zatorre, 2001). The
652 representation of source activity also suggests that the ventral
653 tegmental area is involved during the chills, despite it not
654 being one of the structures we targeted in our analysis. The
655 OFC is involved in reward processing (Berridge, 2003) and is
656 strongly connected with reward structures such as the amygdala,
657 parahippocampal cortex, and medial prefrontal cortex (Bechara
658 et al., 1994). It is implicated in the system underlying autonomic
659 responses to music (Koelsch, 2014), and activation of the right
660 OFC has been correlated with pleasantness ratings (Blood et al.,
661 1999) and intensity of chills (Blood and Zatorre, 2001). It has
662 been established that the intensity of chills is correlated with the
663 intensity of striatum and OFC activity (Blood and Zatorre, 2001;
664 Salimpoor et al., 2011) and that theta activity could be a cortical
665 activity related to reward structure activation. Considering these
666 facts in addition to the gradual increase of theta activity in the
667 prefrontal ROI when the emotional ratings increased, we can
668 reasonably hypothesize that the increased power of the theta
669 activity is linked to the intensity of the felt pleasure. Further
670 theta activity over fronto-central sites might be dependent on
671 arousal level (Sammler et al., 2007). As suggested by works from
672 Jäncke et al. (2015), the increased activity in the right versus left
673 OFC that was observed with source localization corresponded
674 to the significant right prefrontal increased theta activity. It
675 is also one more argument in favor of the implication of the
676 OFC in the production of theta prefrontal activity. However, the
677 same overall trends found in both the left and right prefrontal
678 areas suggest that musical pleasure is reflected on both left
679 and right frontal sites, independent of positive or negative
680 emotional valence.

681 The SMA is known to be associated with the reward system
682 (Knutson et al., 2001), and its involvement has previously
683 been identified during musical chills using PET; the rCBF
684 increased in the SMA during chills (Blood and Zatorre, 2001).



729 Furthermore, a correlation analysis demonstrated that the
730 rCBF in the SMA (as well as the insula and the OFC) was
731 positively correlated with increase in several psychophysiological
732 parameters (HR, respiration rate, and electromyography).
733 Decreased theta oscillatory activity in the central region may be
734 related to the specific SMA activity during chills. Both source
735 localization and surface activity showed a large difference for
736 chills compared with the two other conditions, and there were
737 no differences between low versus high pleasure. At this point,
738 we are not able to explain the stronger effect of theta decrease
739 in the right compared with the left central ROI, even though—
740 as for the prefrontal areas—the same trends were found in both
741 left and right ROIs.

742 The results of our study also highlighted a gradual decrease
743 of theta power for the right temporal ROI and an increased
744 activity in both the left and the right superior temporal gyrus.
745 Because the superior temporal gyrus is known to be activated
746 in the processing of musical liking/disliking (Brown et al., 2004),
747 and familiarity influences neural responses in the auditory cortex
748 (Pereira et al., 2011), we hypothesized that this temporal activity
749 is related to the auditory cortex. The auditory cortex localized in
750 the superior temporal gyrus is connected to the OFC, and these
751 two structures exchange information during music processing
752 (Brown et al., 2004). The activity, which was identified with
753 source localization from the right superior temporal gyrus by
754 Joucla et al. (2018), is likely to be a signature of music liking, and
755
756
757
758
759
760
761
762
763
764
765
766
767
768
769
770
771
772
773
774
775
776
777
778
779
780
781
782
783
784
785
786
787
788
789
790
791
792
793
794
795
796
797
798
799

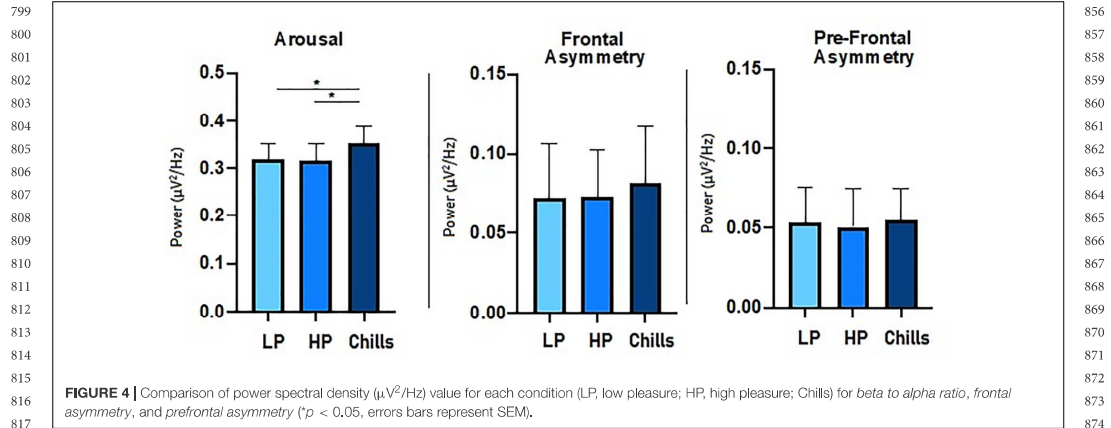


FIGURE 4 | Comparison of power spectral density ($\mu\text{V}^2/\text{Hz}$) value for each condition (LP, low pleasure; HP, high pleasure; Chills) for beta to alpha ratio, frontal asymmetry, and prefrontal asymmetry ($p < 0.05$, errors bars represent SEM).

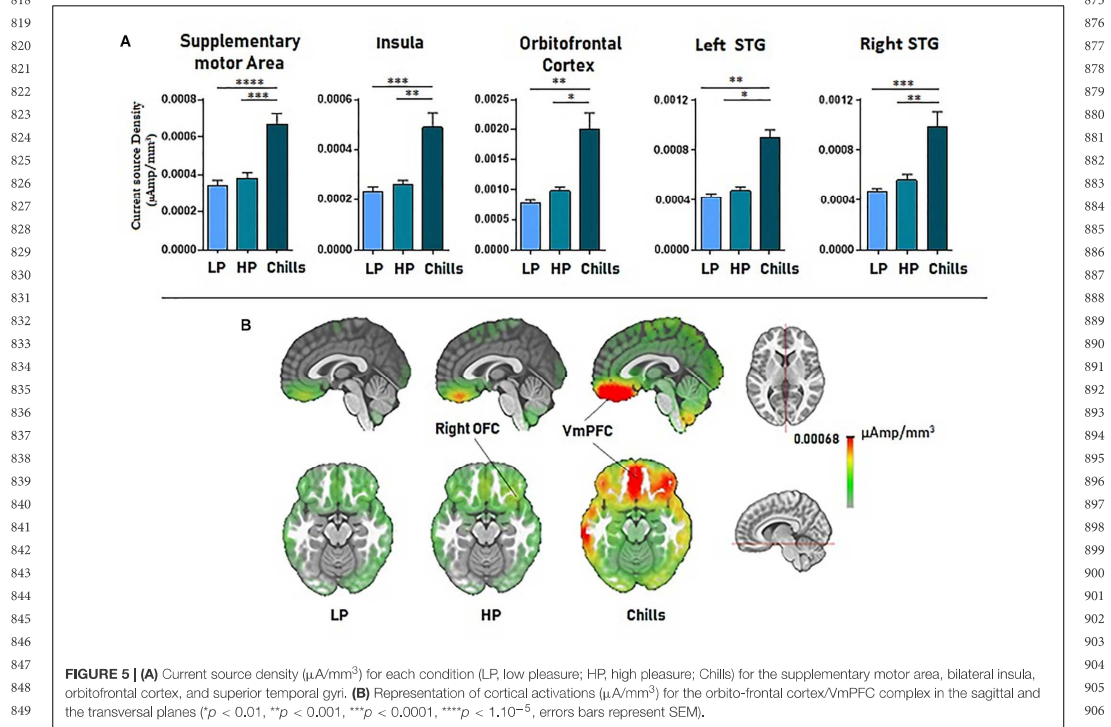


FIGURE 5 | (A) Current source density ($\mu\text{A}/\text{mm}^3$) for each condition (LP, low pleasure; HP, high pleasure; Chills) for the supplementary motor area, bilateral insula, orbitofrontal cortex, and superior temporal gyri. (B) Representation of cortical activations ($\mu\text{A}/\text{mm}^3$) for the orbito-frontal cortex/VmPFC complex in the sagittal and the transversal planes ($p < 0.01$, $**p < 0.001$, $***p < 0.0001$, $****p < 1.10^{-5}$, errors bars represent SEM).

these findings seem to be consistent with ours. Furthermore, a study predicted that people who frequently experience musical chills have a higher structural connectivity (larger volume of white matter connectivity) between the prefrontal medial cortex,

insula, and posterior superior temporal gyri (Sachs et al., 2016). We suggest that the theta activity in the right temporal ROI could be related to music liking processing. However, we observed increased activation of both the bilateral insula and the left and

913 right temporal gyri with higher emotional ratings. It is difficult to
914 draw a clear conclusion about the right temporal surface activity
915 because the insula and the auditory cortex are adjacent structures
916 that could both have been involved.

917 We found no specific effect both for oscillatory rhythms
918 in the alpha band or for alpha frontal and prefrontal
919 asymmetry. We hypothesized that the climax of an emotional
920 increase—related to the activation of the approach dimension
921 relative to positive/rewarding stimulation (Omicie et al.,
922 2015)—from neutral to chills should gradually activate the
923 approach dimension, resulting in increased alpha asymmetry.
924 The extracts for each participant were not restricted to only
925 positive or negative valence, and this should have affected
926 the asymmetry results. The major or minor mode of the
927 music can transmit happy or sad emotions, which, respectively,
928 increase or decrease the alpha power in the left frontal
929 area (Schmidt and Trainor, 2001; Tsang et al., 2001). The
930 experience of chills is quite difficult to reproduce on demand
931 so we did not restrict the valence of the stimuli because this
932 was neither an aim nor would have been a gain for these
933 investigations. We considered that such a restriction was not
934 in line with our perspectives and would have reflected the
935 cerebral activity relative to only positively or negatively valenced
936 pleasurable musical chills.

937 The arousal calculation showed an increased β/α ratio on
938 frontal regions. There was no effect on the alpha band over
939 the frontal region; however, the same non-significant trends
940 can be identified for most of the participants, who exhibited
941 decreased alpha power on the frontal ROI during chills, which
942 explains our results. Several structures were identified using
943 fMRI, such as the amygdala, prefrontal cortex, or even auditory
944 cortex, which were linked with arousal changes while listening
945 to music (Daly et al., 2019). Further connectivity analysis using
946 fMRI has already demonstrated a close relationship between the
947 nucleus accumbens, ventral tegmental area, and hypothalamus
948 in investigations of affective responses to music (Menon and
949 Levitin, 2005). Connectivity analysis also highlighted a close
950 relationship between several structures of the affective and
951 autonomic systems. The hypothalamus and particularly the
952 insula, which showed connectivity with the nucleus accumbens
953 during music listening (Menon and Levitin, 2005), corresponds
954 to the sources localized in our study.

955 There are some limitations to our experiment. Because of
956 experimental constraints, the chills might have been “auto-
957 induced” because participants expected the specific chill-
958 inducing moment that they had previously indicated. This could
959 have reinforced the anticipation phase highlighted by Salimpoor
960 et al. (2011). Thus, the investigation of the chills induced by
961 music in a natural ecological setting could provide somewhat
962 different EEG results. Furthermore, several participants reported
963 that the “aseptic” environment of our experimental room could
964 influence the emotions, and thus the experience of chills. For this
965 reason, some participants did not report enough chills and were
966 therefore excluded from EEG analysis. More occurrences of chills
967 per subject, as well as a larger sample, would have reinforced the
968 overall EEG results and the robustness of effects. The influence of
969 the headphones on the local electric field has not been evaluated

970 and the motor activity related to continuous reports might have
971 slightly influenced the oscillatory activity.

972 CONCLUSION

973 In conclusion, our results suggest that HD-EEG could provide
974 relevant information about musical pleasure and pleasurable
975 musical chills mainly in the right prefrontal, temporal, and
976 central areas in the theta band. This work is a first step for the
977 study of musical chills using EEG in more ecological paradigms.
978 In future experiments, with the development of wireless mobile
979 EEG systems and coupled with physiological measures, EEG
980 recordings characterizing musical pleasure and musical chills
981 in both social contexts and live conditions would provide a
982 relevant parameter to study emotional synchronization of groups
983 (Chabin et al., 2020).

984 DATA AVAILABILITY STATEMENT

985 The datasets presented in this study can be found in online
986 repositories. The names of the repository/repositories and
987 accession number(s) can be found below: [https://figshare.com/
988 articles/Cerebral_activity_during_peak_emotion_in_response_
989 to_music_revealed_by_High-Density_EEG_/11687868](https://figshare.com/articles/Cerebral_activity_during_peak_emotion_in_response_to_music_revealed_by_High-Density_EEG_/11687868).

990 ETHICS STATEMENT

991 The studies involving human participants were reviewed and
992 approved by the CPP Ouest V – Rennes; no. 2018-A01653-
993 52. The patients/participants provided their written informed
994 consent to participate in this study.

1000 AUTHOR CONTRIBUTIONS

1001 TC, AC, DG, and LP: conceptualization. TC, CJ, LM, and
1002 LP: investigations. TC and TC: data curation. TC and DG:
1003 formal analysis. TC, DG, and LP: writing—original draft. TC, AC,
1004 DG, EH, TM, and LP: writing—review and editing. All authors
1005 contributed to the article and approved the submitted version.

1006 FUNDING

1007 This study was supported by a grant from the University Hospital
1008 of Besançon (Appel à Projet Interne CHU – APICHU no.
1009 API/2018/92) and benefited from financial support from the
1010 Clinical and Integrative Laboratory of Neurosciences of Besançon
1011 for device acquisitions. The sponsors played no role in the study.

1012 ACKNOWLEDGMENTS

1013 The authors thank all the volunteers participants included
1014 in the study. The authors also thank Jennifer Dobson for
1015 proofreading the article.


1027 REFERENCES

- 1028 Acquadro, M. A. S., Congedo, M., and De Ridder, D. (2016). Music performance
1029 as an experimental approach to hyperscanning studies. *Front. Hum. Neurosci.*
1030 10:185. doi: 10.3389/fnhum.2016.00242
- 1031 Aftanas, L. I., and Golocheikine, S. A. (2001). Human anterior and frontal
1032 midline theta and lower alpha reflect emotionally positive state and internalized
1033 attention: high-resolution EEG investigation of meditation. *Neurosci. Lett.* 310,
1034 57–60. doi: 10.1016/S0304-3940(01)02094-8
- 1035 Altenmüller, E., Schürmann, K., Lim, V. K., and Parlitz, D. (2002). Hits to the left,
1036 flops to the right: different emotions during listening to music are reflected in
1037 cortical lateralisation patterns. *Neuropsychologia* 40, 2242–2256. doi: 10.1016/
- 1038 s0028-3932(02)00107-0
- 1039 Arjmand, H.-A., Hohagen, J., Paton, B., and Rickard, N. S. (2017). Emotional
1040 responses to music: shifts in frontal brain asymmetry mark periods of musical
1041 change. *Front. Psychol.* 8:2044. doi: 10.3389/fpsyg.2017.02044
- 1042 Bechara, A., Damasio, A. R., Damasio, H., and Anderson, S. W. (1994). Insensitivity
1043 to future consequences following damage to human prefrontal cortex. *Cognition*
1044 50, 7–15. doi: 10.1016/0010-0277(94)90018-3
- 1045 Berridge, K. C. (2003). Pleasures of the brain. *Brain Cogn.* 52, 106–128. doi: 10.
1046 1016/s0278-2626(03)00014-9
- 1047 Bevilacqua, D., Davidesco, I., Wan, L., Chaloner, K., Rowland, J., Ding,
1048 M., et al. (2018). Brain-to-brain synchrony and learning outcomes
1049 vary by student–teacher dynamics: evidence from a real-world
1050 classroom electroencephalography study. *J. Cogn. Neurosci.* 31, 401–411.
1051 doi: 10.1162/jocn_a_01274
- 1052 Blood, A. J., and Zatorre, R. J. (2001). Intensely pleasurable responses to music
1053 correlate with activity in brain regions implicated in reward and emotion. *Proc.*
1054 *Natl. Acad. Sci. U.S.A.* 98, 11818–11823. doi: 10.1073/pnas.191355898
- 1055 Blood, A. J., Zatorre, R. J., Bermudez, P., and Evans, A. C. (1999). Emotional
1056 responses to pleasant and unpleasant music correlate with activity in paralimbic
1057 brain regions. *Nat. Neurosci.* 2, 382–387. doi: 10.1038/7299
- 1058 Brattico, E., and Pearce, M. (2013). The neuroaesthetics of music. *Psychol. Aesth.*
1059 *Creat. Arts* 7, 48–61. doi: 10.1037/a0031624
- 1060 Brown, S., Martinez, M. J., and Parsons, L. M. (2004). Passive music listening
1061 spontaneously engages limbic and paralimbic systems. *Neuroreport* 15, 2033–
1062 2037. doi: 10.1097/00001756-200409150-00008
- 1063 Brunet, D., Murray, M. M., and Michel, C. M. (2011). Spatiotemporal analysis
1064 of multichannel EEG: CARTOOL. *Comput. Intell. Neurosci.* 2011:813870. doi:
1065 10.1155/2011/813870
- 1066 Burgess, A. P., and Gruzelier, J. H. (1997). Short duration synchronization of
1067 human theta rhythm during recognition memory. *Neuroreport* 8, 1039–1042.
1068 doi: 10.1097/00001756-199703030-00044
- 1069 Chabin, T., Tio, G., Comte, A., Joucla, C., Gabriel, D., and Pazart, L. (2020).
1070 The relevance of a conductor competition for the study of emotional
1071 synchronization within and between groups in a natural musical setting. *Front.*
1072 *Psychol.* 10:2954. doi: 10.3389/fpsyg.2019.02954
- 1073 Daly, I., Williams, D., Hwang, F., Kirke, A., Miranda, E. R., and Nasuto, S. J. (2019).
1074 Electroencephalography reflects the activity of sub-cortical brain regions during
1075 approach-withdrawal behaviour while listening to music. *Sci. Rep.* 9, 1–22.
1076 doi: 10.1038/s41598-019-45105-2
- 1077 Davidson, R. J. (2004). What does the prefrontal cortex “do” in affect: perspectives
1078 on frontal EEG asymmetry research. *Biol. Psychol.* 67, 219–233. doi: 10.1016/j.
- 1079 biopsycho.2004.03.008
- 1080 Dikker, S., Wan, L., Davidesco, I., Kaggen, L., Oostrik, M., McClintock, J., et al.
1081 (2017). Brain-to-brain synchrony tracks real-world dynamic group interactions
1082 in the classroom. *Curr. Biol.* 27, 1375–1380. doi: 10.1016/j.cub.2017.04.002
- 1083 Egermann, H., Sutherland, M. E., Grewe, O., Nagel, F., Kopiez, R., and Altenmüller,
1084 E. (2011). Does music listening in a social context alter experience? A
1085 physiological and psychological perspective on emotion. *Music. Sci.* 15, 307–
1086 323. doi: 10.1177/1029864911399497
- 1087 Goupil, L., and Aucouturier, J.-J. (2019). Musical pleasure and musical emotions.
1088 *Proc. Natl. Acad. Sci. U.S.A.* 116, 3364–3366. doi: 10.1073/pnas.1900369116
- 1089 Gruber, M. J., Watrous, A. J., Ekstrom, A. D., Ranganath, C., and Otten, L. J.
1090 (2013). Expected reward modulates encoding-related theta activity before an
1091 event. *NeuroImage* 64, 68–74. doi: 10.1016/j.neuroimage.2012.07.064
- 1092 Harmon-Jones, E., Gable, P. A., and Peterson, C. K. (2010). The role of asymmetric
1093 frontal cortical activity in emotion-related phenomena: a review and update.
1094 *Biol. Psychol.* 84, 451–462. doi: 10.1016/j.biopsycho.2009.08.010
- 1095 Jäncke, L., Kühnis, J., Rogenmoser, L., and Elmer, S. (2015). Time course of
1096 EEG oscillations during repeated listening of a well-known aria. *Front. Hum.*
1097 *Neurosci.* 9:401. doi: 10.3389/fnhum.2015.00401
- 1098 Knutson, B., Fong, G. W., Adams, C. M., Varner, J. L., and Hommer, D. (2001).
1099 Dissociation of reward anticipation and outcome with event-related fMRI.
1100 *Neuroreport* 12, 3683–3687. doi: 10.1097/00001756-200112040-00016
- 1101 Koelsch, S. (2014). Brain correlates of music-evoked emotions. *Nat. Rev. Neurosci.*
1102 15, 170–180. doi: 10.1038/nrn3666
- 1103 Krause, C. M., Pörn, B., Lang, A. H., and Laine, M. (1999). Relative alpha
1104 desynchronization and synchronization during perception of music. *Scand. J.*
1105 *Psychol.* 40, 209–215. doi: 10.1111/1467-9450.00118
- 1106 Lewis, M. (2005). Bridging emotion theory and neurobiology through dynamic
1107 systems modeling. *Behav. Brain Sci.* 28, 169–194; discussion 194. doi: 10.1017/
- 1108 s0140525x0500004x
- 1109 Lin, Y.-P., Duann, J.-R., Chen, J.-H., and Jung, T.-P. (2010).
1110 Electroencephalographic dynamics of musical emotion perception
1111 revealed by independent spectral components. *NeuroReport* 21, 410–415.
1112 doi: 10.1097/WNR.0b013e31833774de
- 1113 Lindenberger, U., Li, S.-C., Gruber, W., and Müller, V. (2009). Brains swinging
1114 in concert: cortical phase synchronization while playing guitar. *BMC Neurosci.*
1115 10:22. doi: 10.1186/1471-2202-10-22
- 1116 Mas-Herrero, E., Marco Pallarés, J., Lorenzo-Seva, U., Zatorre, R., and Rodriguez-
1117 Fornells, A. (2013). Individual differences in music reward experiences. *Music*
1118 *Percept.* 31, 118–138. doi: 10.1525/mp.2013.31.2.118
- 1119 Matusz, P. J., Dikker, S., Huth, A. G., and Perrodin, C. (2019). Are we ready for
1120 real-world neuroscience? *J. Cogn. Neurosci.* 31, 327–338. doi: 10.1162/jocn_e-
- 1121 01276
- 1122 Menendez, R. G., de P., Andino, S. G., Lantz, G., Michel, C. M., and Landis,
1123 T. (2001). Noninvasive localization of electromagnetic epileptic activity. I.
1124 Method descriptions and simulations. *Brain Topogr.* 14, 131–137. doi: 10.1023/
- 1125 A:1012944913650
- 1126 Mennella, R., Patron, E., and Palomba, D. (2017). Frontal alpha asymmetry
1127 neurofeedback for the reduction of negative affect and anxiety. *Behav. Res. Ther.*
1128 92, 32–40. doi: 10.1016/j.brat.2017.02.002
- 1129 Menon, V., and Levitin, D. J. (2005). The rewards of music listening: response and
1130 physiological connectivity of the mesolimbic system. *NeuroImage* 28, 175–184.
1131 doi: 10.1016/j.neuroimage.2005.05.053
- 1132 Michel, C. M., and Brunet, D. (2019). EEG source imaging: a practical review of the
1133 analysis steps. *Front. Neurol.* 10:325. doi: 10.3389/fneur.2019.00325
- 1134 Müller, V., Sängler, J., and Lindenberger, U. (2013). Intra- and inter-brain
1135 synchronization during musical improvisation on the guitar. *PLoS One*
1136 8:e73852. doi: 10.1371/journal.pone.0073852
- 1137 Müller, V., Sängler, J., and Lindenberger, U. (2018). Hyperbrain network properties
1138 of guitarists playing in quartet. *Ann. N. Y. Acad. Sci.* 1423, 198–210. doi: 10.
1139 1111/nyas.13656
- 1140 Nemati, S., Akrami, H., Salehi, S., Esteky, H., and Moghimi, S. (2019). Lost in
1141 music: neural signature of pleasure and its role in modulating attentional
1142 resources. *Brain Res.* 1711, 7–15. doi: 10.1016/j.brainres.2019.01.011
- 1143 Oldfield, R. C. (1971). The assessment and analysis of handedness: the edinburgh
1144 inventory. *Neuropsychologia* 9, 97–113. doi: 10.1016/0028-3932(71)90067-4
- 1145 Omigie, D., Dellacherie, D., Hasboun, D., George, N., Clement, S., Baulac, M., et al.
1146 (2015). An intracranial EEG study of the neural dynamics of musical valence
1147 processing. *Cereb. Cortex* 25, 4038–4047. doi: 10.1093/cercor/bhu118
- 1148 Pereira, C. S., Teixeira, J., Figueiredo, P., Xavier, J., Castro, S. L., and Brattico,
1149 E. (2011). Music and emotions in the brain: familiarity matters. *PLoS One*
1150 6:e27241. doi: 10.1371/journal.pone.0027241
- 1151 Pornpattananangkul, N., and Nusslock, R. (2016). Willing to wait: elevated
1152 reward-processing EEG activity associated with a greater preference for
1153 larger-but-delayed rewards. *Neuropsychologia* 91, 141–162. doi: 10.1016/j.
- 1154 neuropsychologia.2016.07.037
- 1155 Pu, M., and Yu, R. (2019). Post-encoding frontal theta activity predicts incidental
1156 memory in the reward context. *Neurobiol. Learn. Mem.* 158, 14–23. doi: 10.
1157 1016/j.nlm.2019.01.008
- 1158
- 1159
- 1160

- 1141 Ramirez, R., Planas, J., Escude, N., Mercade, J., and Farriols, C. (2018). EEG-
1142 based analysis of the emotional effect of music therapy on palliative care cancer
1143 patients. *Front. Psychol.* 9:254. doi: 10.3389/fpsyg.2018.00254
- 1144 Ramirez, R., and Vamvakousis, Z. (2012). "Detecting emotion from EEG signals
1145 using the emotive Epoc device," in *Brain Informatics*, eds F. M. Zanzotto, S.
1146 Tsumoto, N. Taatgen, and Y. Yao (Berlin: Springer), 175–184. doi: 10.1007/978-
1147 3-642-35139-6_17
- 1148 Ramos, J., and Corsi-Cabrera, M. (1989). Does brain electrical activity react to
1149 music? *Int. J. Neurosci.* 47, 351–357. doi: 10.3109/00207458908987449
- 1150 Rogenmoser, L., Zollinger, N., Elmer, S., and Jäncke, L. (2016). Independent
1151 component processes underlying emotions during natural music listening. *Soc.
1152 Cogn. Affect. Neurosci.* 11, 1428–1439. doi: 10.1093/scan/nsw048
- 1153 Sachs, M. E., Ellis, R. J., Schlaug, G., and Loui, P. (2016). Brain connectivity reflects
1154 human aesthetic responses to music. *Soc. Cogn. Affect. Neurosci.* 11, 884–891.
1155 doi: 10.1093/scan/nsw009
- 1156 Saliba, J., Lorenzo-Seva, U., Marco-Pallares, J., Tillmann, B., Zeitouni, A., and
1157 Lehmann, A. (2016). French validation of the barcelona music reward
1158 questionnaire. *PeerJ* 4:e1760. doi: 10.7717/peerj.1760
- 1159 Salimpoor, V. N., Benovoy, M., Larcher, K., Dagher, A., and Zatorre, R. J. (2011).
1160 Anatomically distinct dopamine release during anticipation and experience of
1161 peak emotion to music. *Nat. Neurosci.* 14, 257–262. doi: 10.1038/nn.2726
- 1162 Salimpoor, V. N., Benovoy, M., Longo, G., Cooperstock, J. R., and Zatorre, R. J.
1163 (2009). The rewarding aspects of music listening are related to degree of
1164 emotional arousal. *PLoS One* 4:e7487. doi: 10.1371/journal.pone.0007487
- 1165 Sammler, D., Grigutsch, M., Fritz, T., and Koelsch, S. (2007). Music and emotion:
1166 electrophysiological correlates of the processing of pleasant and unpleasant
1167 music. *Psychophysiology* 44, 293–304. doi: 10.1111/j.1469-8986.2007.00497.x
- 1168 Sängler, J., Müller, V., and Lindenberger, U. (2013). Directionality in hyperbrain
1169 networks discriminates between leaders and followers in guitar duets. *Front.
1170 Hum. Neurosci.* 7:234. doi: 10.3389/fnhum.2013.00234
- 1171 Schmidt, L., and Trainor, L. (2001). Frontal brain electrical activity (EEG)
1172 distinguishes valence and intensity of musical emotions. *Cogn. Emot.* 15,
1173 487–500. doi: 10.1080/02699930126048
- 1174 Soto, V., Tyson-Carr, J., Kokmotou, K., Roberts, H., Cook, S., Fallon, N., et al.
1175 (2018). Brain responses to emotional faces in natural settings: a wireless mobile
1176 EEG recording study. *Front. Psychol.* 9:2003. doi: 10.3389/fpsyg.2018.02003
- 1177
1178
1179
1180
1181
1182
1183
1184
1185
1186
1187
1188
1189
1190
1191
1192
1193
1194
1195
1196
1197
- Steiger, T. K., and Bunzeck, N. (2017). Reward dependent invigoration relates
1198 to theta oscillations and is predicted by dopaminergic midbrain integrity
1199 in healthy elderly. *Front. Aging Neurosci.* 9:1. doi: 10.3389/fnagi.2017.
1200 00001
- 1201 Swarbrick, D., Bosnyak, D., Livingstone, S. R., Bansal, J., Marsh-Rollo, S.,
1202 Woolhouse, M. H., et al. (2018). How live music moves us: head movement
1203 differences in audiences to live versus recorded music. *Front. Psychol.* 9:2682.
1204 doi: 10.3389/fpsyg.2018.02682
- 1205 Tsang, C. D., Trainor, L. J., Santesso, D. L., Tasker, S. L., and Schmidt, L. A. (2001).
1206 Frontal EEG responses as a function of affective musical features. *Ann. N. Y.
1207 Acad. Sci.* 930, 439–442. doi: 10.1111/j.1749-6632.2001.tb05764.x
- 1208 Vecchiato, G., Maglione, A. G., Scorpecci, A., Malerba, P., Marsella, P., Di
1209 Francesco, G., et al. (2012). "EEG frontal asymmetry related to pleasantness
1210 of music perception in healthy children and cochlear implanted users," in
1211 *Proceedings of the 2012 Annual International Conference of the IEEE Engineering
1212 in Medicine and Biology Society*, San Diego, CA, 4740–4743. doi: 10.1109/
1213 EMBC.2012.6347026
- 1214 Volpe, G., D'Ausilio, A., Badino, L., Camurri, A., and Fadiga, L. (2016). Measuring
1215 social interaction in music ensembles. *Philos. Trans. R. Soc. Lond. Ser. B Biol.
1216 Sci.* 371:1693. doi: 10.1098/rstb.2015.0377
- 1217 Wingerden, M., van Vinck, M., Lankelma, J., and Pennartz, C. M. A.
1218 (2010). Theta-band phase locking of orbitofrontal neurons during reward
1219 expectancy. *J. Neurosci.* 30, 7078–7087. doi: 10.1523/JNEUROSCI.3860-09.
1220 2010
- Conflict of Interest:** The authors declare that the research was conducted in the
1221 absence of any commercial or financial relationships that could be construed as a
1222 potential conflict of interest.
- 1223
1224
1225
1226
1227
1228
1229
1230
1231
1232
1233
1234
1235
1236
1237
1238
1239
1240
1241
1242
1243
1244
1245
1246
1247
1248
1249
1250
1251
1252
1253
1254

Article

Behavioral and Electrophysiological Arguments in Favor of a Relationship between Impulsivity, Risk-Taking, and Success on the Iowa Gambling Task

Julie Giustiniani ^{1,2,3,*} , Coralie Joucla ², Djamila Bennabi ^{1,2,4}, Magali Nicolier ^{1,2,3}, Thibault Chabin ², Caroline Masse ^{1,2}, Benoît Trojak ^{5,6}, Pierre Vandel ^{1,2,3}, Emmanuel Haffen ^{1,2,3,4} and Damien Gabriel ^{2,3}

¹ Department of Clinical Psychiatry, University Hospital of Besançon, 25000 Besançon, France; djamilabennabi@univ-fcomte.fr (D.B.); mnicolier@chu-besancon.fr (M.N.); cmasse@chu-besancon.fr (C.M.); pierre.vandel@univ-fcomte.fr (P.V.); emmanuel.haffen@univ-fcomte.fr (E.H.)

² Laboratory of Neurosciences, University of Burgundy Franche-Comté, 25000 Besançon, France;

coralie.joucla.pro@gmail.com (C.J.); tchabin@edu.univ-fcomte.fr (T.C.); dgabriel@chu-besancon.fr (D.G.)

³ Clinical Investigation Centre, University Hospital of Besançon, Inserm, CIC 1431, 25000 Besançon, France

⁴ FondaMental Foundation, 94000 Créteil, France

⁵ Department of Psychiatry and Addictology, University Hospital of Dijon, 21079 Dijon, France; benoit.trojak@chu-dijon.fr

⁶ EA 4452, LPPM, University of Burgundy Franche-Comté, 21000 Dijon, France

* Correspondence: jgiustiniani@chu-besancon.fr or julie.giustiniani@univ-fcomte.fr; Tel.: +33-3-8121-8154

Received: 29 July 2019; Accepted: 21 September 2019; Published: 24 September 2019



Abstract: The aim of the present study was to investigate the relationship between trait impulsivity, risk-taking, and decision-making performance. We recruited 20 healthy participants who performed the Iowa Gambling Task (IGT) and the Balloon Analog Risk Task (BART) to measure decision-making and risk-taking. The impulsivity was measured by the Barratt Impulsiveness Scale. Resting-state neural activity was recorded to explore whether brain oscillatory rhythms provide important information about the dispositional trait of impulsivity. We found a significant correlation between the ability to develop a successful strategy and the propensity to take more risks in the first trials of the BART. Risk-taking was negatively correlated with cognitive impulsivity in participants who were unable to develop a successful strategy. Neither risk-taking nor decision-making was correlated with cortical asymmetry. In a more exploratory approach, the group was sub-divided in function of participants' performances at the IGT. We found that the group who developed a successful strategy at the IGT was more prone to risk, whereas the group who failed showed a greater cognitive impulsivity. These results emphasize the need for individuals to explore their environment to develop a successful strategy in uncertain situations, which may not be possible without taking risks.

Keywords: decision-making; IGT; BART; impulsivity; risk-taking; theta oscillations

1. Introduction

Decision-making is a complex cognitive process that is used, in part, to solve open and risky problems in daily life, the outcomes of which are often unpredictable [1,2]. Impairments in decision-making can have harmful direct consequences on the social and personal aspects of one's daily life. Such impairment has been observed in many neuropsychiatric disorders [3–5], such as behavioral addictions [6] and substance addictions [7–9].

The Iowa Gambling Task (IGT) is a psychological task designed to assess impairments in decision-making under conditions of uncertainty by simulating real-life economic decisions [10].

The IGT is principally used to describe neurophysiological and clinical differences between pathological and healthy populations [7–11], however, within-group differences are frequent and expected on this task. Indeed, several studies have reported a failure rate of up to 55% on the IGT in healthy populations for both the original and adapted versions [7,12–15].

If the IGT was originally designed to support the somatic marker hypothesis [1,16], several other factors appear to have a major impact on decision-making. Since the IGT was founded on the principle that real-life decision-making involves choices that are based on expected but uncertain rewards and penalties, and that optimal choices are based on well-considered strategies, one can suggest that impulse control problems may alter our decisions. Impulsivity appears to be one factor that could explain the differences in IGT performance. Impulsivity is considered an important element in the decision-making process; however, it is a complex construct with several components, making the discernment of its exact role difficult. Impulsivity is nevertheless known to be related to risk-taking and a lack of reflection between environmental stimuli and the behavioral response during decision-making [17]. Impulsivity is often measured by self-report, such as the Barratt Impulsiveness Scale (BIS) [18]. Self-report scales can provide valuable information about impulsivity as a stable trait in individuals [19]. Impulsivity is a trait known for being able to influence decision-making performance [20,21], however, the influence of impulsive personality traits on IGT performance remains unclear [22,23]. Although self-report scales can measure relatively stable characteristics, they are subjective and cannot directly measure the biological foundations of impulsivity. To provide a more objective and biologically based measure, the Balloon Analogue Risk Task (BART) was specifically designed to measure the propensity for risk-taking [24,25]. Impulsivity is indeed closely related to risk-taking. High levels of impulsivity often result in risky behaviors, such as the initiation of alcohol or drug use [26]. At the same time, almost any human decision carries some risk. Decision-making and risk are two inter-related factors, both being related to various uncertainties [27]. Tests of performance on the IGT and BART in subjects with high and low trait impulsivity revealed that trait impulsivity did not lead to different scores on the IGT or BART. Nevertheless, participants with low trait impulsivity were seemingly more risk averse on the BART, although they made more advantageous selections at the IGT [28]. Clarifying how these concepts are entangled appears to be important in both theoretical and clinical aspects. Changes in decisions are frequently justified by an important feeling, led by impulsivity or risk-taking with a lack of distinction of how each of these constructs specifically influences behaviors and decisions. Indeed, the somatic marker hypothesis refers to the collection of signals related to the body and the brain that characterize the emotional and affective responses [29]. In this sense, the impact of impulsivity or risk-taking on decision-making does not contradict the somatic marker hypothesis.

In this context, the aim of the present study was to further investigate the relationship between impulsivity, risk-taking, and IGT performance. We examined whether the ability to develop a correct strategy on the IGT requires specific levels of trait impulsivity and risk-taking on the BART. To provide important additional information about impulsivity as a dispositional trait, we also recorded neural activity at rest. With resting-state electroencephalography (EEG), tonic cortical activity prior to task performance can be measured, reflecting individual behavioral and cognitive disposition, which has shown to be stable for several years [30]. Therefore, baseline anterior asymmetries can be used as a predictor of performance on behavioral tasks [31]. The left prefrontal cortex (PFC) was previously associated to individual predisposition to respond in terms of approach-related and reward-related motivations and emotions [32–35]. Based on these previous results, we hypothesized that hemispheric asymmetry predicts behavioral profiles with regard to performance on the BART, potentially explaining IGT performance. Exploring such a neural and behavioral relationship may have direct implications on the understanding of mechanisms underlying many neuropsychiatric disorders in which decision-making is altered.

2. Materials and Methods

2.1. Participants

Twenty healthy, right-handed subjects (10 male and 10 female, mean age = 38.7 years, SD = 18.3 years, range = 21–59 years) participated in the study. Participants had no previous medical history of psychiatric disorders, substance abuse, alcohol abuse, neurological disease, traumatic brain injury, or stroke, and did not take any medication.

The participants received information regarding the aims and procedures of the study and were provided with written informed consent to participate. Because of the influence of real money on motivation, the subjects were informed that the monetary payment would be proportional to their gains in the game, with an exchange rate of 1%. Because of ethical considerations, regardless of their performance, all of the participants received 75€ at the end of the study. The protocol was approved by the Committee of Protection of Persons (CPP-Est II) and was conducted in accordance with the principles of the Declaration of Helsinki (ClinicalTrials.gov identifier: NCT02873572) (for details, see Giustiniani et al., 2015 [14]).

2.2. Experimental Tasks

2.2.1. Iowa Gambling Task

We used an electronic version of the IGT [10], which was adapted to study event-related potentials and to analyze brain activity. The aim of the IGT was to win as much money as possible by making successive selections between four decks [10,36].

A full description of the study is presented elsewhere [14]. Minimal changes were made to adapt the IGT to EEG. First, to extend electrophysiological recordings of the exploration phase, the number of trials was increased from 100 to 200 and the subjects were given no hints about the presence of advantageous or disadvantageous decks. Second, the design of the trials was modified to minimize ocular artifacts. In each trial, subjects were instructed to fix their gaze on a cross or letter while making their selection by pressing a key. After the selection, feedback regarding the deck that was chosen and total credit amount was displayed, followed by the amount of money won in the trial. A fixation point then appeared to orient the eyes to the center of the screen, followed by a fixed letter that announced the result.

Once the task was complete, the subjects' performance and psychometric test response were analyzed (for details, see Giustiniani et al., 2015 [14]), and sub-groups were constituted. The sub-groups did not differ with regard to age ($p = 0.21$), level of education ($p = 0.62$), or sex ($p = 0.37$). Additionally, no differences were found between groups in various psychometric tests that evaluated neuropsychiatric disorders, addiction, or personality dimensions. Namely, there were no differences on the South Oaks Gambling Screen (SOGS) [37], Alcohol Use Disorder Identification Test AUDIT [38], Diminuer Entourage Trop Cannabis/Cut-down Annoyed Guilty Eye-opener DETC/CAGE [38], Fagerström test [38], Beck Depression Inventory, abbreviated version (BDI) [39], Liebowitz scale [40], or the Big Five Inventory-French (BFI-Fr) [41] (see Table 1). Sub-groups were analyzed in a more exploratory approach, to get more insight into the mechanisms observed on the whole group of participants.

Table 1. Psychometric data. Self-report scales were used to measure pathological gambling with the South Oaks Gambling Screen (SOGS) [37], alcohol dependence with the Alcohol Use Disorder Identification Test (AUDIT) [38], the noxious use of cannabis with the Diminuer Entourage Trop Cannabis/Cut-down Annoyed Guilty Eye-opener DETC/CAGE [38], addiction to nicotine with the Fagerström test [38], the existence and intensity of depressive symptoms with the Beck Depression Inventory abbreviated version (BDI) [39], anxiety with the Liebowitz scale [40], impulsivity with the Barratt Impulsiveness Scale version 10 (BIS-10) [42], and individual differences in the five personality dimensions with the Big Five Inventory-French (BFI-Fr) [41].

	Favorable Groups	Undecided Group	Favorable versus Undecided Group	
	Average (SD)	Average (SD)	t-Value	p-Value (Two Tailed)
BDI	1.11 (0.93)	2 (2.14)	-1.15	0.26
BFI-Fr				
Neuroticism	21.33 (5.36)	18.63 (7.03)	0.95	0.36
Extraversion	27.33 (5.70)	26.09 (6.11)	0.47	0.65
Openness to experience	36.33 (6.12)	35.54 (5.64)	0.29	0.77
Agreeableness	41.55 (5.08)	42.18 (6.29)	-0.24	0.81
Conscientiousness	33.67 (7.50)	36.45 (4.99)	-0.99	0.33
LIEBOWITZ SCALE				
Performance	6.36 (1.49)	5.20 (1.08)	0.64	0.53
Anxiety	6.05 (1.96)	5.91 (0.77)	0.08	0.93
Avoidance	7.44 (4.27)	6.45 (5.14)	0.46	0.65
Social interaction	4.67 (3.74)	5.36 (4.13)	-0.39	0.7
Anxiety	6.67 (1.57)	4.50 (0.96)	1.06	0.3
Avoidance	7.78 (5.47)	5.18 (4.64)	1.15	0.27
AUDIT	5.55 (5.34)	3.81 (4.24)	0.81	0.43
Fagerström Scale	4.55 (3.74)	2.36 (2.16)	1.64	0.12
DETC/CAGE	0.44 (1.33)	0.18 (0.60)	0.59	0.56
SOGS	0	0	-	-
	0	0.09 (0.30)	-0.9	0.38

2.2.2. EEG Recording at Rest

An EEG recording at rest was performed immediately after the IGT. During the recording, subjects were instructed to close their eyes and remain still. Resting-state EEGs were recorded for four minutes using OSG digital equipment (BrainRT, OSG bvba, Rumst, Belgium) with two Schwarzer AHNS epas 44-channel amplifiers (Natus, Munich, Germany). A 64-channel electrode cap (Easycap, easycapGmbH, Ammersee, Germany) with electrodes positioned in the 10/10 system was used to get the EEG signals. EEG data were continuously recorded with a band pass of 0.05–100 Hz and sampling rate of 1000 Hz. All electrode impedances were below 30 k Ω before recording was started.

Two electrodes (AF3 and AF4) were placed over the left and right dorsolateral PFC. Signal processing was performed using EEGLab and Cartool software (<http://brainmapping.unige.ch/Cartool.php>). Raw EEG data were digitally low-pass filtered and corrected for eye movements. Raw EEG data were offline re-referenced to a common average. Channels with excessive noise (due to malfunctioning or bad signal during data collection) were replaced using topographic interpolation. EEG signals that contained residual muscle movements or other forms of artifacts at >50 μ V were removed prior to further analysis. Each artifact-free period was further segmented into epochs of 2 s. A fast-Fourier transform (Hamming window: length 10%) was used to estimate spectral power in the θ (4–7 Hz) and α (8–12 Hz) frequency bands. Finally, frontal brain asymmetry was calculated by subtracting the mean power densities (AF3–AF4) in each frequency band.

2.2.3. Balloon Analog Risk Task

The BART is an informatic-based measure of the propensity for risk-taking [24]. During the BART, the participants were required to sequentially press a button to inflate a series of 30 balloons that were displayed on the computer screen. The computer screen showed a small simulated balloon accompanied by a balloon pump, a reset button labeled “Collect \$\$\$,” and a permanent display of

money earned, labeled “Total Earned”. Each pump corresponded to 5 cents that accrued in a temporary reserve (the amount of money in this reserve was never indicated to the participant). At any point during each balloon inflation, the participant could stop pumping the balloon and click the “Collect \$\$\$” button. Clicking this button transferred all of the money from the temporary reserve to the permanent bank, at which time the new total that was earned was incrementally updated cent-by-cent while a slot machine payoff sound effect played. The balloon could either grow larger or explode. A larger balloon was simultaneously associated with a greater probability of explosion and a larger virtual monetary reward. Once a balloon was pumped past its individual explosion point, it exploded with a “pop” sound, at which time all of the money in the temporary reserve was then lost. After each balloon explosion or money collection, the participant’s exposure to that balloon ended and a new balloon appeared, until all 30 balloons had been displayed. All of the balloons had a different explosion threshold. Balloon breakpoints ranged from 1 to 128 pumps. The participants were instructed to maximize the amount of virtual reward during the experiment. Risk-taking behavior on the BART was measured by calculating the total and mean adjusted number of pumps across balloons (only trials in which the balloons did not explode were included in the calculations).

2.3. Data Analyses

To compare the IGT with other parameters (trait impulsivity, BART scores, and frontal asymmetry), the individual net score was calculated by subtracting the number of disadvantageous decks from the number of advantageous decks that were selected in the last 60 trials. We chose not to focus on the total net score because some studies have proposed that performance at the IGT is more reliable when participants have developed a certain strategy [43]. The last 60 trials was the chosen cutoff point at which we considered that participants had started developing a strategy because the net score for the last 60 trials was significantly higher than the net score in the first trials (see Giustiniani et al., 2015 for more details [14]). The subjects were classified into two groups according to their net score on the IGT. Based on classifications set in previous studies [7,10,29], nine subjects were classified as favorable (net score > 10) and 11 subjects were classified as undecided (net score = 10 to –10). No subject had a net score less than –10.

Trait impulsivity was measured by the BIS-10 [18,42] and included total and subscale scores (cognitive, motor, and non-planning). There were two parameters of the BART considered for analysis, namely total number of adjusted pumps, which was directly related to money earned at the end of the experiment (each adjusted pump corresponded to 5 cents earned), and average number of adjusted pumps, which was calculated by dividing the number of completed trials by the total number of adjusted pumps. As a greater average suggests a greater inclination of participants to take risks, the second parameter is more related to risk-taking.

The performance on the BART as it pertains to trials was taken into account by analyzing the task in three blocks with 10 trials each [24,25]. To determine if data from the favorable and undecided groups had a common mean during throughout their participation in the BART, repeated-measures analysis of variance (ANOVA) were performed. The threshold of significance was set to 5% and post hoc analyses were performed using Fischer’s Least Significant Difference (LSD) test.

Relationships between net score on the IGT, trait impulsivity, risk-taking, and frontal asymmetry were assessed using nonparametric Spearman rank-order correlations. To consider multiple comparisons, the threshold of significance was set to 1%. We performed the analysis using Statistica 11.0 for Windows (StatSoft, Inc., Tulsa, OK).

Given significant relationships for the participants as a whole, additional correlations were then evaluated specifically in the favorable and undecided groups.

3. Results

3.1. Relationship between the Participants' Strategy on the IGT and Risk-Taking on the BART

We investigated whether risk-taking would influence performances on the IGT by studying the relationship between IGT scores and presses on the BART. We observed that both the average adjusted number of pumps and the total number of pumps were reduced in the first ten trials of the BART compared to the other blocks. A repeated-measures analysis of variance (ANOVA), with block (trials 1–10, trials 11–20, trials 21–30) as factor, confirmed these observations by showing significant differences on the average adjusted number of pumps during the different blocks ($F_{2,38} = 10.9982$, $p < 0.001$), with reduced performances in trials 1–10 compared to 11–20 ($p < 0.05$, Bonferroni corrected) and 21–30 ($p < 0.001$, Bonferroni corrected).

To study whether the reduced performances in the first 10 trials could be imputable to the development of an efficient strategy, we explored their relationship with IGT performances. The first trials of the BART display an uncertainty level which confers a structure close to the last trials of the IGT [44]. A significant correlation was found between the net score on the IGT during the last 60 trials and the total adjusted number of pumps in the first 10 trials of the BART (Spearman $R = 0.48$, $t_{18} = 2.38$, $p = 0.03$; Figure 1). No such relationship was found between the total net score nor the net score during the first 40 trials and the average adjusted number of pumps. Furthermore, no relationship was found between the IGT net score during the last 60 trials of the IGT and the total adjusted number of pumps in the trials 11–20 (Spearman $R = -0.03$, $t_{18} = -0.15$, $p = 0.88$) and 21–30 (Spearman $R = -0.02$, $t_{18} = -0.07$, $p = 0.94$) of the BART.

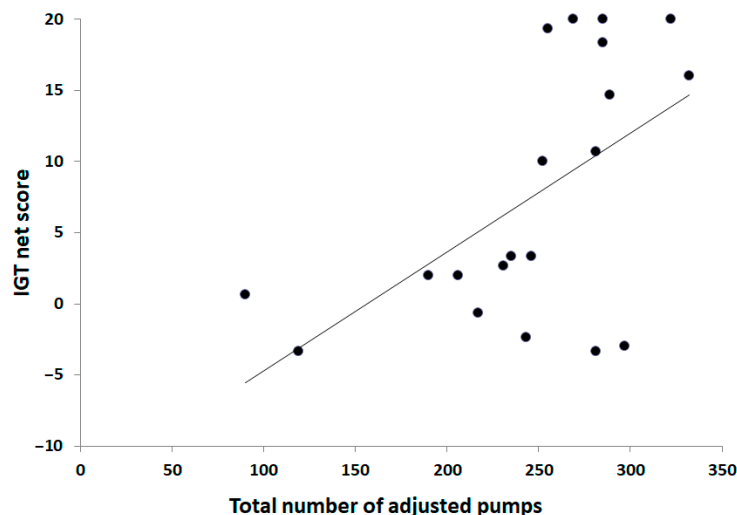


Figure 1. Scatterplot of the relationship between the total number of adjusted pumps in the 10 first trials and the net score on the Iowa Gambling Task (IGT) during the last 60 trials for all participants.

To explore further the observations made on the whole group of participants, we investigated whether risk-taking at the BART was more pronounced in participants able to develop a successful strategy. The data were analyzed using repeated-measures ANOVA, with block (trials 1–10, trials 11–20, trials 21–30) and group (favorable, undecided) as factors. A significant block \times group interaction ($F_{2,36} = 6.2811$, $p < 0.01$; Figure 2) was observed. The group who failed to develop a favorable strategy showed less risk-taking, i.e. fewer presses compared to the Favorable group in the first 10 trials of the

BART (Fisher LSD test, $p < 0.05$). Risk-taking of the undecided group was also reduced in the first 10 trials compared to trials 11–20 and trials 21–30 ($p < 0.01$ for both comparisons). No relationship between the BART and the IGT was found on the average number of adjusted responses.

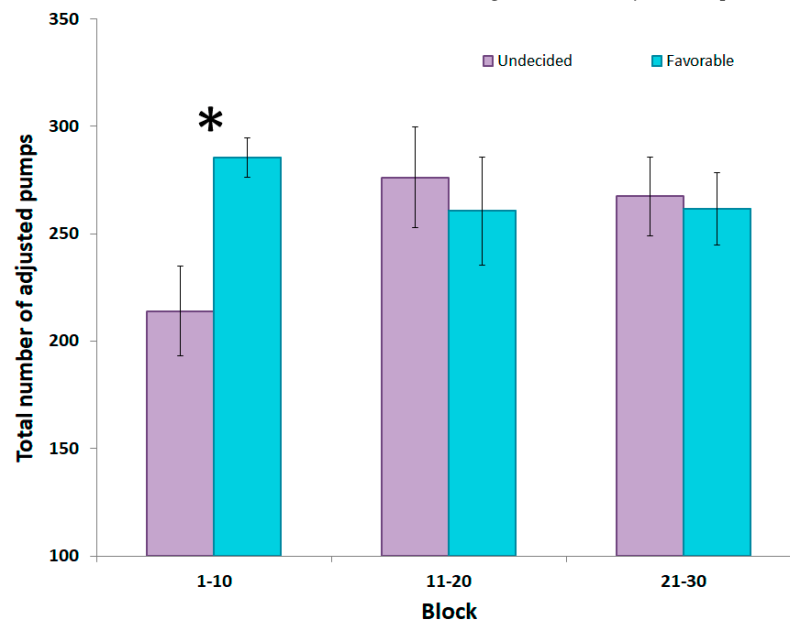


Figure 2. Mean total number of adjusted pumps across the three blocks of the Balloon Analog Risk Task (BART) in the undecided and favorable groups. A significant difference was found between the groups in the first 10 trials. * $p < 0.05$.

3.2. Relationship between the Participants' Strategy on the IGT and Impulsivity

No relationship was found between net score on the IGT and BIS-10 global (Spearman $R = -0.08$, $t_{18} = -0.35$, $p = 0.73$) as well as the subscale scores. This absence of relationship could also be observed for the favorable and the undecided groups.

3.3. Relationship between Impulsivity and Risk-Taking

We also evaluated whether risk-taking could be defined as a component of impulsivity by studying the relationship between impulsivity scales and risk-taking (Table 2). Surprisingly, there was a significant negative correlation between the average number of adjusted pumps on the BART and cognitive impulsivity on the BIS-10 in the whole group (Spearman $R = -0.68$, $t_{18} = -3.89$, $p < 0.01$), suggesting that cognitive impulsivity is associated with a decreased propensity in risk-taking. This negative correlation particularly strong in the undecided group (Spearman $R = -0.77$, $t_9 = -3.59$, $p < 0.01$) but not in the favorable group ($p > 0.05$). No correlation was found between the total number of adjusted pumps on the BART and the global BIS score and all subscales (all $ps > 0.05$).

Table 2. Spearman's correlation between BART scores (average and total number of adjusted pumps) and (Barratt Impulsiveness Scale version 10) BIS-10 scores. Spearman R values are followed by t values in brackets. * $p < 0.01$. When significance was obtained for the overall group of participants, additional significance was analyzed for the favorable and undecided groups.

	BART					
	Whole Subjects $N = 20$	Average		Whole Subjects $N = 20$	Total	
		Advantageous $N = 9$	Undecided $N = 11$		Advantageous $N = 9$	Undecided $N = 11$
BIS-10 motor	-0.16 (-0.69)	-	-	0.39 (1.80)	-	-
BIS-10 cognitive	-0.68 (-3.89) *	0.16 (0.44)	-0.68 (-3.59) *	-0.33 (-1.50)	-	-
BIS-10 no planning	-0.35 (-1.58)	-	-	-0.27 (-1.20)	-	-
BIS-10 total	-0.50 (-2.43)	-	-	-0.03 (-0.14)	-	-

3.4. Relationship between Cortical Asymmetry and Behavioral Scores

Frontal asymmetry was measured to examine how the dispositional trait of impulsivity could affect risk-taking behavior. No significant relationship was observed between IGT net scores and theta asymmetry ($p = 0.81$) or alpha asymmetry ($p = 0.12$). There was also no relationship between BART, theta (total number of pumps: $p = 0.83$; average number of pumps: $p = 0.16$) and alpha asymmetry (total number of pumps: $p = 0.37$; average number of pumps: $p = 0.42$).

We nevertheless questioned whether a subgroup of subjects could be affected by cortical asymmetry. Nonparametric Spearman rank-order correlation analysis on sub-groups revealed a very strong, significant relationship between the mean adjusted number of balloon pumps and theta asymmetry in favorable subjects (Spearman $R = 0.9$, $t_7 = 5.462793$, $p < 0.001$). Figure 3 shows the relationship between frontal theta asymmetry and risk-taking behavior. No significant relationship was found for either the alpha frequency range or undecided subjects. Additionally, no significant relationship was found between the BIS-10 global and subscale scores and theta asymmetry (all $ps > 0.1$).

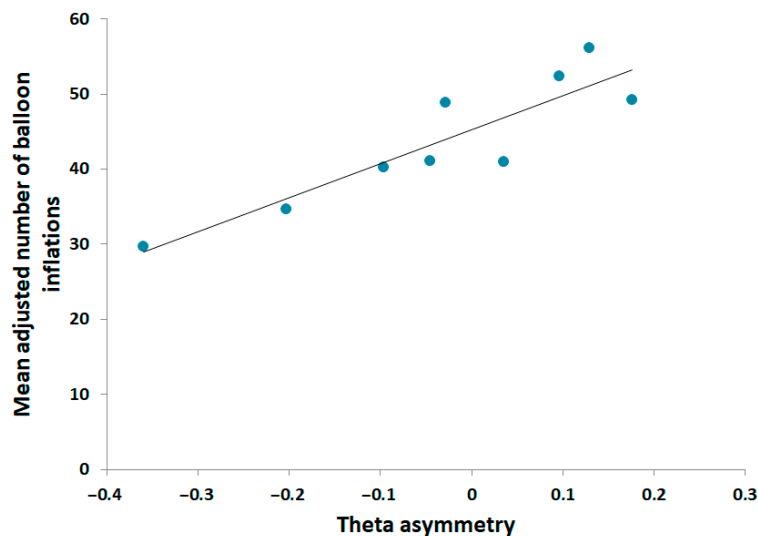


Figure 3. Scatterplot of the relationship between the mean adjusted number of pumps and resting state theta asymmetry in favorable subjects.

We investigated whether participants who were able to develop the correct strategy on the IGT presented differences in impulsivity compared with undecided subjects. Favorable and undecided participants did not differ in any components of the BIS-10 (all p s > 0.1 for global BIS score and all subscales). No significant relationship was found between BIS-10 global and subscale scores and theta asymmetry ($p > 0.1$).

4. Discussion

The present study explored the role of impulsivity and risk-taking on decision-making performance on the IGT and the impact of frontal lobe asymmetry on decision-making.

4.1. Association between Decision-Making, Risk-Taking, and Impulsivity

Although the sensitivity of the IGT in detecting impairments in decision-making under conditions of uncertainty is well established and numerous studies have highlighted the complexity of this task and challenge this pretense for understanding the functions or dysfunctions it purportedly measures [16,45,46]. A positive correlation was found between the net score on the IGT during the conceptual phase (the last 60 trials) and the propensity to inflate the balloon during the first block of the BART. Differences in impulsivity and risk-taking were found between participants who were able to develop a correct strategy and those who could not; indeed, we found that the undecided group made significantly fewer pumps than the favorable group in the first block of the BART. Differences in impulsivity and risk-taking were found between participants who were able to develop a correct strategy and those who could not. The undecided group made significantly fewer pumps than the favorable group in the first block of the BART. The undecided group was more cautious than the favorable group at the beginning of the task, even if both groups had the same level of performances in the following blocks. Compared to the undecided group, favorable group took more risk at the beginning of the task, to achieve more rapidly a high level of performance.

To understand this association between the later stage of the IGT and BART scores, risk-taking in the early and late stages of the IGT should be considered separately [45]. This is different from the BART, which was specifically designed to measure the propensity for risk-taking from the beginning of the task [24,25]. Nevertheless, the correlation observed between the IGT and the first block of the BART is not inconsistent with the literature. During the BART, participants did not know about the probability distribution of explosions which confers an uncertainty level at the task in addition to the risk, mostly on the first trials [44]. In this context, the first trials of the BART share similar mechanisms as the last trials of the IGT. In contrast to the early stages of the IGT, during which risk-taking is not a deliberate act but rather reflects a failure to recognize risk, risks are explicit in the later stages of the IGT. Thus, the propensity for risk-taking can be measured only at the end of the IGT and only if the subject develops explicit knowledge of the risk profile [28].

Participants who make advantageous decisions in the later trials of the IGT are those who have developed explicit knowledge of the risky decks and the development of such knowledge may be more related to the propensity for risk-taking than trait impulsivity [28]. Our initial hypothesis that we would observe increased risk-taking in the undecided group was not confirmed. It was actually quite the opposite that was observed, the undecided group showing a reduced risk propensity at the beginning of the BART. The subjects in the undecided group were unable to develop the appropriate strategies and consequently did not consciously take risks [28]. Low impulsivity has been previously shown to be correlated with better performance on the IGT and fewer pumps on the BART [28]. Converse to the previous study, in which the groups were formed based on the subjects' impulsivity, group assignment in the present study was based on IGT performance. Furthermore, we did not find a link between the net score on the IGT and impulsivity. This suggests that impulsivity alone cannot explain decision-making abilities, suggesting a complex process. The difference observed in the total number of adjusted pumps in the first block suggests that the favorable group took more risk at the beginning of the BART. These risks accelerated the learning process, thereby allowing the favorable

group to earn more money at the start of the game. Since subjects of the undecided group showed a reduced propensity to take risk in the different tasks, they consequently learned more slowly the probability of explosions at the BART and decks composition at the IGT. More risk-taking is therefore not necessarily related to a pathological process and, in this case, a lower level of risk-taking could lead to bad decision-making. Impact of risk-taking during the IGT appears to be more complex than only unidirectional.

4.2. Association between Impulsivity and Risk-Taking

We found a strong, negative relationship between a high impulsivity level as rated by the cognitive subscale of the BIS-10 and risk-taking. When we continued analysis on sub-group, we found this correlation in the undecided group, but not in the favorable group. The cognitive subscale is related to attentional impulsiveness (i.e., difficulty maintaining attention to relevant details) [47]. Although no correlation was found between impulsivity and IGT performance, the association between impulsivity and risk-taking in the undecided group suggests an indirect link between cognitive impulsivity and the inability to develop an appropriate strategy on the IGT. The difficulty maintaining attention to relevant details may explain why these subjects were unable to develop a correct strategy on the IGT. The lack of a significant relationship between impulsivity on the BIS-10 and risk-taking on the BART in the favorable group does not necessarily imply that there are no impulsive components to this group. The impulsivity construct proposed by Dickman [48] and the two types of impulsivity: Functional and dysfunctional [49] may help explain why differences were found only in the first block of the BART in the group who developed a correct strategy. Functional impulsivity is related to a tendency to make decisions quickly and is principally relevant in situations with personal gain, whereas dysfunctional impulsivity is related to a lack of reflection, resulting in negative consequences. In the first block of the BART, the subjects in the favorable group presented a stronger propensity to inflate. In the other blocks, the subjects in the undecided group tended to pump more to further inflate the balloon, as opposed to subjects in the favorable group, who adjusted their strategy as a function of the risk they perceived. Subjects with the individual trait of functional impulsivity tended to make impulsive decisions when they perceived that strategy high level of risky behavior would result in positive outcomes [48]. Moreover, functional impulsivity has been associated with Gray's concept of reward reactivity [50]. Functional impulsivity is positively related to measures of the Behavioral Approach System (BAS) and negatively related to measures of the Behavioral Inhibition System (BIS). Consistent with this concept, subjects with high levels of functional impulsivity appeared to be more motivated by and reactive to situations that likely lead to reward [50]. Risk-taking has often been associated with motivational drives and behavior [51], measured by a high score on the BAS [52]. Here, we propose that risk-taking is impacted by both impulsivity and motivation, because participants are willing to take risks for the potential of the associated reward. Increased motivation in the favorable group may explain why their behavioral performances were improved following each task. The subjects in the favorable group may have exhibited functional impulsivity and greater motivation to explore the game, therefore leading to development of the correct strategy.

4.3. Association between Neural Activity and Behavioral Performance

The neural response at rest supports the positive influence of risk-taking on behavior to a certain extent. Consistent with a previous study, we found that individual differences in risky choice behavior could be predicted by the hemispheric balance of resting-state PFC activity [30,53,54]. Higher theta-band power detected by the left electrodes was associated with a greater mean adjusted number of pumps in the favorable group. A negative correlation was previously reported between resting-state theta oscillations and glucose metabolism, in which theta oscillations at rest indicated a low level of neural activity [54]. Right theta oscillations may reflect right hypoactivity, which has been associated with a lack of regulatory ability to suppress a more seductive choice, due to immediate higher payoffs [30]. The right PFC is involved in withdrawal-related emotions [55,56] and a logical assumption is that

hypoactivity in the right PFC is negatively correlated with BIS scores. We may postulate that left hemisphere hypoactivity reflects activity of the right hemisphere. Thus, the favorable group should have had better regulatory abilities to suppress immediate appealing options. Thus, subjects who can develop a successful strategy may, therefore, be likely to have lesser left prefrontal activity, favoring right hemisphere activity, which may indicate a better ability of this group to avoid more immediately attractive options with negative consequences in the long term.

An alternative hypothesis is that theta oscillations are also involved in learning and emotion regulation and may be involved in the encoding of information, particularly during active exploratory phases [57]. This specific theta activity shows that the quality of emotional regulation impacts performance. Theta activity has been associated with emotion regulation and encoding information, with both elements being particularly important for the somatic marker hypothesis, which states that emotions affect information encoding and the decision-making process. According to theories of emotion and motivation, the right PFC is related to negative emotions and drives withdrawal behavior, whereas the left PFC is related to positive emotions and drives approach behavior [34,58]. Prior studies have demonstrated a negative correlation with BIS scores and right theta oscillations [54]. This cortical asymmetry is a well-known correlate of approach motivation and risk-taking [52]. Somatic markers traduced an unconscious and conscious knowledge of the risk. This is reflected by a change in the autonomic arousal before decks selections [1,7]. However, the risk anticipation does not seem to imply strategy differences and differences in risk-taking in healthy subjects [14]. Furthermore, risk-taking requires a conscious or unconscious knowledge of the risk level [44]. In the present study, the participants in the favorable group exhibited a dominant approach behavior and consequently took more risks, resulting in positive consequences for their decision-making under conditions of uncertainty. Finally, absence of a significant correlation between theta asymmetry and net score on the IGT confirms that this theta asymmetry exclusively reflected risk-taking and that understanding the results of IGT performances cannot only be explained by differing levels of risk-taking or impulsivity.

4.4. Impulsivity, Risk-Taking, and Decision-Making

Risk-taking and decision-making under conditions of uncertainty are two behaviors in which the likelihood of outcomes is uncertain, but these differ with regard to decisions about probability [59]. Uncertainty and risky decision are related but distinct concept [44]. Indeed, under risky conditions, subjects can use probability models to evaluate the level of danger, reflected by an increase in the number of pumps on the BART. In the beginning stages of the IGT, information on probability is unknown and unusable until the development of specific knowledge of the IGT. The distinction between risk-taking and impulsivity is also difficult to assess [51]. The link between impulsivity that is measured by psychometric tests and BART scores is still debated [48]. The present study showed that the BART may have several indicators of risk-taking (total score and average score), which is one of the elements that comprises impulsivity. Contrary to some observations, risk-taking as evaluated by the BART should not be considered synonymous with impulsivity. Since a risk-taking is not synonymous of pathological process, several aspects of behavior must be taken into account before hastily concluding a pathological trait in a specific population.

4.5. Limitations and Recommendations for Future Works

Although there are considerable conclusions to be considered from this study, limitations must be noted. The main limitation of the present study is the relatively small number of volunteers in our subgroups classified according to their net score at the IGT. In that respect, results on subgroups are exploratory and toned to be confirmed on a larger scale.

Future works should also take into account any possible age and gender differences in decision-making and risk management, which were not considered in our analysis. It is unclear whether gender may have a repercussion on the results of the current study, as men and women tend to perform at the same level on the IGT, although women tend to take additional trials before

selecting long-term advantageous decks [60]. However, it is worth note that women tend to take less risks than men do in a wide range of behaviors [61]. Age is also an important parameter to take into account, as older adults tend to make more risky decisions relative to younger adults on the IGT, but are more risk averse relative to younger adults on the BART [62]. Future studies should consider how decision-making, risk-taking and impulsivity interact in a pathological population to more clearly define each pathological behavior.

5. Conclusions

In conclusion, we propose that the development of correct decision-making under uncertain conditions is mediated by individuals' predisposition to perceive and orient themselves with their emotions and ability to encode information and use it to explore. In uncertain situations, risk-taking is sometimes necessary to explore the environment and different options that are available. Right/left theta asymmetry may reflect an individual predisposition to be motivated and quickly explore strategies to achieve better performance at the beginning of the BART.

Author Contributions: The following authors conceived of and designed the experiments: J.G., D.G., M.N., P.V., and E.H.; performed the experiments: J.G. and D.G.; analyzed the data: J.G., D.G., C.J., and T.C.; contributed reagents/materials/analysis tools: J.G., D.G., M.N., P.V. and E.H.; contributed to writing the manuscript: J.G., D.G., D.B., M.N., B.T., C.M. and E.H.

Funding: This study was supported by a grant from the French Ministry of Health. (Programme Hospitalier de Recherche Clinique no. 2009-A00806-51). The sponsor played no role in the study.

Conflicts of Interest: The authors confirm that they have no conflict of interest.

References

1. Bechara, A.; Damasio, A.R. The somatic marker hypothesis: A neural theory of economic decision. *Games Econ. Behav.* **2005**, *52*, 336–372. [[CrossRef](#)]
2. Platt, M.L.; Huettel, S.A. Risky business: the neuroeconomics of decision making under uncertainty. *Nat. Neurosci.* **2008**, *11*, 398–403. [[CrossRef](#)] [[PubMed](#)]
3. Cavadini, P.; Riboldi, G.; D'Annunzi, A.; Belotti, P.; Cisima, M.; Bellodi, L. Decision-making heterogeneity in obsessive-compulsive disorder: Ventromedial prefrontal cortex function predicts different treatment outcomes. *Neuropsychologia* **2002**, *40*, 205–211. [[CrossRef](#)]
4. Cavadini, P.; Bassi, T.; Ubbiali, A.; Casolari, A.; Giordani, S.; Zorzi, C.; Bellodi, L. Neuropsychological investigation of decision-making in anorexia nervosa. *Psychiatry Res.* **2004**, *127*, 259–266. [[CrossRef](#)] [[PubMed](#)]
5. Jollant, F.; Guillaume, S.; Jaussent, I.; Castelnaud, D.; Malafosse, A.; Courtet, P. Impaired decision-making in suicide attempters may increase the risk of problems in affective relationships. *J. Affect. Disord.* **2007**, *99*, 59–62. [[CrossRef](#)] [[PubMed](#)]
6. Cavadini, P.; Riboldi, G.; Keller, R.; D'Annunzi, A.; Bellodi, L. Frontal lobe dysfunction in pathological gambling patients. *Biol. Psychiatry* **2002**, *51*, 334–341. [[CrossRef](#)]
7. Bechara, A.; Damasio, H. Decision-making and addiction (part I): impaired activation of somatic states in substance dependent individuals when pondering decisions with negative future consequences. *Neuropsychologia* **2002**, *40*, 1675–1689. [[CrossRef](#)]
8. Bechara, A.; Dolan, S.; Hindes, A. Decision-making and addiction (part II): myopia for the future or hypersensitivity to reward? *Neuropsychologia* **2002**, *40*, 1690–1705. [[CrossRef](#)]
9. Tanabe, J.; Thompson, L.; Claus, E.; Dalwani, M.; Hutchison, K.; Banich, M.T. Prefrontal cortex activity is reduced in gambling and nongambling substance users during decision-making. *Hum. Brain Mapp.* **2007**, *28*, 1276–1286. [[CrossRef](#)]
10. Bechara, A.; Damasio, A.R.; Damasio, H.; Anderson, S.W. Insensitivity to future consequences following damage to human prefrontal cortex. *Cognition* **1994**, *50*, 7–15. [[CrossRef](#)]
11. Oberg, S.A.K.; Christie, G.J.; Tata, M.S. Problem gamblers exhibit reward hypersensitivity in medial frontal cortex during gambling. *Neuropsychologia* **2011**, *49*, 3768–3775. [[CrossRef](#)] [[PubMed](#)]

12. Bagnoux, V.; Thomassin, N.; Gonthier, C.; Roulin, J.-L. Working memory in the processing of the Iowa Gambling Task: an individual differences approach. *PLoS ONE* **2013**, *8*, e81498. [[CrossRef](#)] [[PubMed](#)]
13. Bechara, A.; Dolan, S.; Denburg, N.; Hindes, A.; Anderson, S.W.; Nathan, P.E. Decision-making deficits, linked to a dysfunctional ventromedial prefrontal cortex, revealed in alcohol and stimulant abusers. *Neuropsychologia* **2001**, *39*, 376–389. [[CrossRef](#)]
14. Giustiniani, J.; Gabriel, D.; Nicolier, M.; Monnin, J.; Haffen, E. Neural Correlates of Successful and Unsuccessful Strategic Mechanisms Involved in Uncertain Decision-Making. *PLoS ONE* **2015**, *10*, e0130871. [[CrossRef](#)] [[PubMed](#)]
15. Mapelli, D.; Di Rosa, E.; Cavalletti, M.; Schiff, S.; Tamburin, S. Decision and dopaminergic system: an ERPs study of Iowa gambling task in Parkinson's disease. *Front. Psychol.* **2014**, *5*, 684. [[CrossRef](#)]
16. Dunn, B.D.; Dalgleish, T.; Lawrence, A.D. The somatic marker hypothesis: a critical evaluation. *Neurosci. Biobehav. Rev.* **2006**, *30*, 239–271. [[CrossRef](#)]
17. Xu, S.; Korczykowski, M.; Zhu, S.; Rao, H. Assessment of risk-taking and impulsive behaviors: A comparison between three tasks. *Soc. Behav. Personal.* **2013**, *41*, 477–486. [[CrossRef](#)]
18. Patton, J.H.; Stanford, M.S.; Barratt, E.S. Factor structure of the Barratt impulsiveness scale. *J. Clin. Psychol.* **1995**, *51*, 768–774. [[CrossRef](#)]
19. Swann, A.C.; Bjork, J.M.; Moeller, F.G.; Dougherty, D.M. Two models of impulsivity: relationship to personality traits and psychopathology. *Biol. Psychiatry* **2002**, *51*, 988–994. [[CrossRef](#)]
20. Grassi, G.; Pallanti, S.; Righi, L.; Figeo, M.; Mantione, M.; Denys, D.; Piccagliani, D.; Rossi, A.; Stratta, P. Think twice: Impulsivity and decision making in obsessive-compulsive disorder. *J. Behav. Addict.* **2015**, *4*, 263–272. [[CrossRef](#)]
21. Ochoa, C.; Alvarez-Moya, E.M.; Penelo, E.; Aymami, M.N.; Gómez-Peña, M.; Fernández-Aranda, F.; Granero, R.; Vallejo-Ruiloba, J.; Menchón, J.M.; Lawrence, N.S.; et al. Decision-making deficits in pathological gambling: the role of executive functions, explicit knowledge and impulsivity in relation to decisions made under ambiguity and risk. *Am. J. Addict.* **2013**, *22*, 492–499. [[CrossRef](#)]
22. Goudriaan, A.E.; Oosterlaan, J.; de Beurs, E.; van den Brink, W. Decision making in pathological gambling: a comparison between pathological gamblers, alcohol dependents, persons with Tourette syndrome, and normal controls. *Brain Res. Cogn. Brain Res.* **2005**, *23*, 137–151. [[CrossRef](#)] [[PubMed](#)]
23. Franken, I.H.A.; van Strien, J.W.; Nijs, I.; Muris, P. Impulsivity is associated with behavioral decision-making deficits. *Psychiatry Res.* **2008**, *158*, 155–163. [[CrossRef](#)] [[PubMed](#)]
24. Lejuez, C.W.; Read, J.P.; Kahler, C.W.; Richards, J.B.; Ramsey, S.E.; Stuart, G.L.; Strong, D.R.; Brown, R.A. Evaluation of a behavioral measure of risk taking: The Balloon Analogue Risk Task (BART). *J. Exp. Psychol. Appl.* **2002**, *8*, 75–84. [[CrossRef](#)] [[PubMed](#)]
25. Lejuez, C.W.; Aklin, W.M.; Zvolensky, M.J.; Pedulla, C.M. Evaluation of the Balloon Analogue Risk Task (BART) as a predictor of adolescent real-world risk-taking behaviours. *J. Adolesc.* **2003**, *26*, 475–479. [[CrossRef](#)]
26. Ríos-Bedoya, C.F.; Wilcox, H.C.; Piazza, M.; Anthony, J.C. Children taking risks: the association with cocaine and other drug use by young adulthood. *Addict. Behav.* **2008**, *33*, 1154–1161. [[CrossRef](#)] [[PubMed](#)]
27. Lu, J.; Jain, L.C.; Zhang, G. Risk Management in Decision Making. In *Handbook on Decision Making: Vol 2: Risk Management in Decision Making*; Lu, J., Jain, L.C., Zhang, G., Eds.; Intelligent Systems Reference Library; Springer: Berlin/Heidelberg, Germany, 2012; pp. 3–7, ISBN 978-3-642-25755-1.
28. Upton, D.J.; Bishara, A.J.; Ahn, W.-Y.; Stout, J.C. Propensity for risk taking and trait impulsivity in the Iowa Gambling Task. *Personal. Individ. Differ.* **2011**, *50*, 492–495. [[CrossRef](#)] [[PubMed](#)]
29. Bechara, A.; Damasio, H.; Damasio, A.R. Emotion, decision making and the orbitofrontal cortex. *Cereb. Cortex N. Y. N 1991* **2000**, *10*, 295–307. [[CrossRef](#)] [[PubMed](#)]
30. Gianotti, L.R.R.; Knoch, D.; Faber, P.L.; Lehmann, D.; Pascual-Marqui, R.D.; Diezi, C.; Schoch, C.; Eisenegger, C.; Fehr, E. Tonic activity level in the right prefrontal cortex predicts individuals' risk taking. *Psychol. Sci.* **2009**, *20*, 33–38. [[CrossRef](#)]
31. Schutter, D.J.L.G.; de Haan, E.H.F.; van Honk, J. Anterior asymmetrical alpha activity predicts Iowa gambling performance: distinctly but reversed. *Neuropsychologia* **2004**, *42*, 939–943. [[CrossRef](#)]
32. Balconi, M.; Finocchiaro, R. Decisional impairments in cocaine addiction, reward bias, and cortical oscillation “unbalance”. *Neuropsychiatr. Dis. Treat.* **2015**, *11*, 777–786. [[CrossRef](#)] [[PubMed](#)]

33. Balconi, M.; Mazza, G. Brain oscillations and BIS/BAS (behavioral inhibition/activation system) effects on processing masked emotional cues. ERS/ERD and coherence measures of alpha band. *Int. J. Psychophysiol. Off. J. Int. Organ. Psychophysiol.* **2009**, *74*, 158–165.
34. Davidson, R.J. What does the prefrontal cortex “do” in affect: perspectives on frontal EEG asymmetry research. *Biol. Psychol.* **2004**, *67*, 219–233. [[CrossRef](#)] [[PubMed](#)]
35. Harmon-Jones, E. Contributions from research on anger and cognitive dissonance to understanding the motivational functions of asymmetrical frontal brain activity. *Biol. Psychol.* **2004**, *67*, 51–76. [[CrossRef](#)] [[PubMed](#)]
36. Bechara, A.; Damasio, H.; Tranel, D.; Damasio, A.R. Deciding advantageously before knowing the advantageous strategy. *Science* **1997**, *275*, 1293–1295. [[CrossRef](#)] [[PubMed](#)]
37. Lejoyeux, M. *Abrégés Addictologie*; Masson: Paris, France, 2009.
38. Reynaud, M. *Usage nocif de substances psychoactives*; La documentation française: Paris, France, 2002.
39. Beck, A.T.; Steer, R.A.; Carbin, M.G. Psychometric properties of the Beck Depression Inventory: Twenty-five years of evaluation. *Clin. Psychol. Rev.* **1988**, *8*, 77–100. [[CrossRef](#)]
40. Yao, S.N.; Note, I.; Fanget, F.; Albuissou, E.; Bouvard, M.; Jalenques, I.; Cottraux, J. [Social anxiety in patients with social phobia: Validation of the Liebowitz social anxiety scale: The French version]. *L'Encéphale* **1999**, *25*, 429–435. [[PubMed](#)]
41. Plaisant, O.; Courtois, R.; Réveillère, C.; Mendelsohn, G.A.; John, O.P. Validation par analyse factorielle du Big Five Inventory français (BFI-Fr). Analyse convergente avec le NEO-PI-R. *Ann. Méd.-Psychol. Rev. Psychiatr.* **2010**, *168*, 97–106. [[CrossRef](#)]
42. Bouvard, M. *Questionnaires et échelles d'évaluation de la personnalité*; Elsevier Masson: Paris, France, 2009.
43. Monterosso, J.; Ehrman, R.; Napier, K.L.; O'Brien, C.P.; Childress, A.R. Three decision-making tasks in cocaine-dependent patients: Do they measure the same construct? *Addict. Abingdon Engl.* **2001**, *96*, 1825–1837. [[CrossRef](#)]
44. De Groot, K.; Thuri, R. Disentangling Risk and Uncertainty: When Risk-Taking Measures Are Not About Risk. *Front. Psychol.* **2018**, *9*. [[CrossRef](#)]
45. Brand, M.; Recknor, E.C.; Grabenhorst, F.; Bechara, A. Decisions under ambiguity and decisions under risk: Correlations with executive functions and comparisons of two different gambling tasks with implicit and explicit rules. *J. Clin. Exp. Neuropsychol.* **2007**, *29*, 86–99. [[CrossRef](#)] [[PubMed](#)]
46. Buelow, M.T.; Suhr, J.A. Construct validity of the Iowa Gambling Task. *Neuropsychol. Rev.* **2009**, *19*, 102–114. [[CrossRef](#)] [[PubMed](#)]
47. Gomide Vasconcelos, A.; Sergeant, J.; Corrêa, H.; Mattos, P.; Malloy-Diniz, L. When self-report diverges from performance: The usage of BIS-11 along with neuropsychological tests. *Psychiatry Res.* **2014**, *218*, 236–243. [[CrossRef](#)] [[PubMed](#)]
48. Vigil-Colet, A. Impulsivity and decision making in the balloon analogue risk-taking task. *Personal. Individ. Differ.* **2007**, *43*, 37–45. [[CrossRef](#)]
49. Dickman, S.J. Functional and dysfunctional impulsivity: Personality and cognitive correlates. *J. Pers. Soc. Psychol.* **1990**, *58*, 95–102. [[CrossRef](#)] [[PubMed](#)]
50. Smillie, L.D.; Jackson, C.J. Functional impulsivity and reinforcement sensitivity theory. *J. Pers.* **2006**, *74*, 47–84. [[CrossRef](#)] [[PubMed](#)]
51. Panwar, K.; Rutherford, H.J.V.; Mencl, W.E.; Lacadie, C.M.; Potenza, M.N.; Mayes, L.C. Differential associations between impulsivity and risk-taking and brain activations underlying working memory in adolescents. *Addict. Behav.* **2014**, *39*, 1606–1621. [[CrossRef](#)] [[PubMed](#)]
52. Black, C.L.; Goldstein, K.E.; LaBelle, D.R.; Brown, C.W.; Harmon-Jones, E.; Abramson, L.Y.; Alloy, L.B. Behavioral approach system sensitivity and risk taking interact to predict left-frontal EEG asymmetry. *Behav. Ther.* **2014**, *45*, 640–650. [[CrossRef](#)] [[PubMed](#)]
53. Schutter, D.J.L.G.; Van Honk, J. Electrophysiological ratio markers for the balance between reward and punishment. *Brain Res. Cogn. Brain Res.* **2005**, *24*, 685–690. [[CrossRef](#)]
54. Studer, B.; Pedroni, A.; Rieskamp, J. Predicting Risk-Taking Behavior from Prefrontal Resting-State Activity and Personality. *PLoS ONE* **2013**, *8*, e76861. [[CrossRef](#)]
55. Sutton, S.K.; Davidson, R.J. Prefrontal Brain Asymmetry: A Biological Substrate of the Behavioral Approach and Inhibition Systems. *Psychol. Sci.* **1997**, *8*, 204–210. [[CrossRef](#)]

56. Harmon-Jones, E.; Allen, J.J. Behavioral activation sensitivity and resting frontal EEG asymmetry: Covariation of putative indicators related to risk for mood disorders. *J. Abnorm. Psychol.* **1997**, *106*, 159–163. [[CrossRef](#)] [[PubMed](#)]
57. Knyazev, G.G. Motivation, emotion, and their inhibitory control mirrored in brain oscillations. *Neurosci. Biobehav. Rev.* **2007**, *31*, 377–395. [[CrossRef](#)] [[PubMed](#)]
58. Harmon-Jones, E. Early Career Award. Clarifying the emotive functions of asymmetrical frontal cortical activity. *Psychophysiology* **2003**, *40*, 838–848. [[CrossRef](#)] [[PubMed](#)]
59. Wang, L.; Zheng, J.; Huang, S.; Sun, H. P300 and Decision Making under Risk and Ambiguity. *Comput. Intell. Neurosci.* **2015**, *2015*. [[CrossRef](#)] [[PubMed](#)]
60. Van den Bos, R.; Homberg, J.; de Visser, L. A critical review of sex differences in decision-making tasks: Focus on the Iowa Gambling Task. *Behav. Brain Res.* **2013**, *238*, 95–108. [[CrossRef](#)]
61. Lighthall, N.R.; Mather, M.; Gorlick, M.A. Acute stress increases sex differences in risk seeking in the balloon analogue risk task. *PLoS ONE* **2009**, *4*, e6002. [[CrossRef](#)]
62. Mata, R.; Josef, A.K.; Samanez-Larkin, G.R.; Hertwig, R. Age differences in risky choice: A meta-analysis. *Ann. N. Y. Acad. Sci.* **2011**, *1235*, 18–29. [[CrossRef](#)] [[PubMed](#)]



© 2019 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<http://creativecommons.org/licenses/by/4.0/>).



ELSEVIER

www.elsevier.com/locate/euroneuro


Modelling decision-making under uncertainty: A direct comparison study between human and mouse gambling data

Lidia Cabeza^a, Julie Giustiniani^{a,b,c}, Thibault Chabin^a,
Bahrie Ramadan^a, Coralie Joucla^a, Magali Nicolier^{a,b,c},
Lionel Pazart^c, Emmanuel Haffen^{a,b,c}, Dominique Fellmann^a,
Damien Gabriel^{a,c}, Yvan Peterschmitt^{a,*}

^aEA-481 Laboratoire de Neurosciences Intégratives et Cliniques de Besançon, Université Bourgogne - Franche-Comté, France

^bDepartment of Clinical Psychiatry, University Hospital, Besançon, France

^cCIC-1431 Inserm, University Hospital, Besançon, France

Received 11 March 2019; received in revised form 15 November 2019; accepted 20 November 2019

Available online xxx

KEYWORDS

Decision-making;
Iowa Gambling Task;
Mouse gambling task

Abstract

Decision-making is a conserved evolutionary process enabling us to choose one option among several alternatives, and relies on reward and cognitive control systems. The Iowa Gambling Task allows the assessment of human decision-making under uncertainty by presenting four card decks with various cost-benefit probabilities. Participants seek to maximise their monetary gain by developing long-term optimal-choice strategies. Animal versions have been adapted with nutritional rewards, but interspecies data comparisons are scarce. Our study directly compares the non-pathological decision-making performance between humans and wild-type C57BL/6 mice. Human participants completed an electronic Iowa Gambling Task version, while mice a maze-based adaptation with four arms baited in a probabilistic way. Our data shows closely matching performance between both species with similar patterns of choice behaviours. However, mice showed a faster learning rate than humans. Moreover, both populations were clustered into good, intermediate and poor decision-making categories with similar proportions. Remarkably, mice characterised as good decision-makers behaved the same as humans of the same category, but slight differences among species are evident for the other two

* Corresponding author.

E-mail address: yvan.peterschmitt@univ-fcomte.fr (Y. Peterschmitt).

<https://doi.org/10.1016/j.euroneuro.2019.11.005>
0924-977X/© 2019 Elsevier B.V. and ECNP. All rights reserved.

Please cite this article as: L. Cabeza, J. Giustiniani and T. Chabin et al., Modelling decision-making under uncertainty: A direct comparison study between human and mouse gambling data, European Neuropsychopharmacology, <https://doi.org/10.1016/j.euroneuro.2019.11.005>

subpopulations. Overall, our direct comparative study confirms the good face validity of the rodent gambling task. Extended behavioural characterisation and pathological animal models should help strengthen its construct validity and disentangle the determinants in animals and humans decision-making.

© 2019 Elsevier B.V. and ECNP. All rights reserved.

1. Introduction

Do animals gamble for food just like humans gamble for money? Most species, including human beings, have to make accurate assessments of the cost-benefit trade-off in reward-driven contexts. For example, the longer an animal is searching for food, the more likely it is to accumulate more and better rewards, however, the more likely it is to encounter a predator. Thus, decision-making (DM) is an essential mechanism for survival in partly predictable situations (de Froment et al., 2014). The underlying principles of how living beings achieve efficient DM are still not fully understood, and assessing human DM using animal models requires further validation. Laboratory tests simulating real-life DM, such as gambling tasks, have been developed. Particularly, the Iowa Gambling Task (IGT) is widely used to assess human DM under uncertainty. Participants are required to maximise the monetary gains by selecting cards from four decks with various cost-benefit probabilities (Bechara et al., 1994). Players face a choice-conflict between playing from long-term disadvantageous decks yielding higher gains but frequent losses, or from advantageous ones associated with smaller immediate rewards and less frequent penalties. Thus, optimal decisions involve the animal refraining from the preferred large immediate rewards whilst opting for lower but more frequent gains, leading to long-term benefits.

The IGT was originally created to study DM impairments in patients with ventromedial prefrontal cortex damage. Compared to healthy participants who progressively learn to select the advantageous decks, the impaired patients decisively underperform by being unable to set the optimal strategy from repeated negative outcomes. Besides this, many IGT studies with healthy populations have shown a high interindividual performance variability (Bechara and Damasio, 2002; Bechara et al., 2002; Steingrover et al., 2013). Indeed, several clinical reports indicated that, while a majority of healthy participants develop the optimal strategy (good DM individuals), others do not acquire a preference for one deck over the others, which is indicative of a lack of learning (poor DM individuals) (Bechara et al., 2001, 2002; Glicksohn et al., 2007a). The literature has shown that many factors may account for this variability in gambling performance. For example, Bechara et al. explain the poor IGT performance in terms of atypical sensitivity to the reward or punishment. Mechanisms underlying this atypical sensitivity are numerous and still not fully elucidated. Among the potential contributors to IGT failure, the presence of personality characteristics associated with risky behaviour (Buelow and Suhr, 2009), of weaknesses in executive functioning (Brand et al., 2007), a low motivation (Giustiniani et al., 2015), a negative affect at the time of task (Suhr and Tsanadis, 2007) as well as low educational

(Davis et al., 2008) and low intellectual levels (Barry and Petry, 2008) have been reported. These data illustrate a behavioural continuum with an overlap in choice strategies between human non-pathological and pathological conditions.

Interestingly, animal versions of the IGT with species-appropriate adjustments to assess DM in a design comparable to humans (van den Bos et al., 2006; van Enkhuizen et al., 2013; Heilbronner, 2017; Pittaras et al., 2016; Rivalan et al., 2009, 2013; de Visser et al., 2011; Young et al., 2011; Zeeb et al., 2009) have been developed. For example, in rodent gambling tasks the decks of cards have been replaced by mazes or operant chambers equipped with different options of varying outcomes; monetary gains are replaced with nutritional rewards since money and food are thought to drive similar behaviours (Lehner et al., 2017).

Similar to human studies, rodents with lesions perform sub-optimally in adapted IGT versions and show a delay in adopting a consistent strategy compared to healthy individuals (Rivalan et al., 2011; Winstanley and Clark, 2015). For the majority of healthy rodents, as the task progresses a raise in advantageous choices is observed, which is commonly accompanied by individuals clustering in three subpopulations who reflect the variable choice strategies (van Enkhuizen et al., 2014; Pittaras et al., 2016; Rivalan et al., 2009). Good DM individuals quickly develop a strong preference for the advantageous options, while poor DM individuals display the worst performances, either not showing any preference for neither advantageous nor disadvantageous options or displaying a long-term preference for the disadvantageous ones. Intermediate DM individuals perform between the other two subgroups. Whether these interindividual variabilities observed in both healthy human and animal populations are similar has never been directly assessed to substantiate the face validity of the models.

Animal and human versions of gambling tasks usually evaluate performance with different analytical methods (Bechara et al., 2001; Zeeb and Winstanley, 2013) and are hence not compared in a direct manner even if animal gambling tasks have been specially created to simulate human DM processes. However, direct comparative research will be a valuable approach allowing clinicians and researchers to bridge their findings, rendering translational studies as a direct tool genuinely meaningful (van Enkhuizen et al., 2014).

Here, we bridge the clinical and preclinical findings on DM processes featuring uncertainty by directly comparing the performance of healthy humans and mice for the first time in the same study using IGT-adapted tasks and identical analytical methods.

Based on the recent literature addressing the validity of mouse gambling tasks (mGT) (van den Bos et al., 2013; van Enkhuizen et al., 2014; Heilbronner, 2017; Pittaras et al., 2016; Winstanley and Clark, 2015), we

forecasted a good face validity of our animal model with a similar overall performance regarding the human population (Giustiniani et al., 2015). To check for the existence of a similar variability in the choice strategies among the species, we then compared the stratification of both populations according to the endpoint performances. To evaluate the construct validity, we further studied the behavioural indexes of the cognitive processes sub-serving the optimal choices. Finally, we sought to find the correlations with an endpoint performance, focusing on the parameters of rigidity (Pittaras, 2013), flexibility (van Enkhuizen et al., 2014) and win-stay and lose-shift choices (de Visser et al., 2011). We also analysed the psychometric scores from human participants and reward sensitivity data in mice, to relate the performance or choice strategies to additional behavioural traits.

2. Experimental procedures

2.1. Participants

To reduce conceptual and methodological differences between the tasks, we controlled factors known to interfere with the results such as sex differences and the presence of instructions, especially in humans. In that respect, we recruited only male subjects because they have been described as less risky than females, choosing the advantageous options more frequently in the IGT (Singh, 2016) and its rodent adaptations (van den Bos et al., 2006). Further characterisations will include both genders, allowing us to take into account the variability we accepted to overlook in the present study (Prendergast et al., 2014).

2.2. Humans

Forty healthy right-handed participants (mean age \pm SEM = 24.7 ± 5.1 ; range 19-38) were involved in the study. None reported a previous medical history of psychiatric disorders, substance or alcohol abuse, neurological diseases, traumatic brain injury or stroke, and none reported taking any medication.

Participants received information regarding the aim of the task and gave their written informed consent to take part in the study. Given the influence of real money playing a significant role on motivation, subjects were informed that the monetary payment would be proportional to the global gain obtained in the task (Meyer, 2004; Meyer et al., 2000; Miedl et al., 2010). Due to the ethical considerations and independent of individual performance, all participants received the maximum amount of €85 at the end of the experiment. The protocol was approved by the Committee of Protection of Persons (CPP-Est-11; authorisation given by the General Health Administration (ANSM 2016-A00870-51 and NCT 02862821)).

2.3. Mice

Forty C57BL/6J mice (Ets Janvier Labs, Saint-Berthevin, France) were used for this study. Inbred animals such as C57BL/6J are known to behave uniformly and display low interindividual variability (Festing, 2014) (but see recent work by Tuttle et al., 2018). All animals were 3-5 months old at the time of testing, group-housed and maintained under a 12 h-circadian cycle with a constant temperature (22 ± 2 °C). Water was available ad libitum and all mice were food-restricted at 80-90% of their free-feeding weight (mean weight (g) \pm SEM = 22.2 ± 0.2) in order to motivate

exploration and completion of the nutritionally rewarding task (Rivalan et al., 2009). Experiments were performed in behavioural rooms with tight luminous intensity. All procedures met the NIH guidelines for the care and use of laboratory animals and were approved by the University of Franche-Comte Animal Care and Use Committee (CEBEA-58). All efforts were taken to minimise animal suffering during the testing according to the Directive from the European Council on the 22nd of September 2010 (2010/63/EU).

2.4. Human gambling task

The task was an adapted electronic version of the IGT (Giustiniani et al., 2015), whose aim was to maximise monetary gain through successive selections between four decks. The decks' composition, values and scheduled reward-penalty were predetermined identically to match the original IGT (Bechara, 1997; Bechara and Damasio, 2005; Bechara et al., 1994). The decks looked identical, but differed in composition. Decks A and B were disadvantageous, yielding immediate rewards but involved major economic losses on the long run. Decks C and D were advantageous, yielding frequent small wins and smaller long-term penalties, resulting in long-term gain. The decks also varied in their schedule of losses, with decks B and D featuring infrequent, decks A and C frequent losses, respectively. All decks contained an infinite number of cards in this computerised version of the IGT. The monetary reward was converted from US Dollars to Euros in order to match the currency of the test subjects. At the beginning of the task, participants had a loan of €2000.

Contrary to most IGT experiments, participants were not provided with information regarding the presence of advantageous or disadvantageous decks and the number of trials. This ensured that a somewhat partial advantage compared to the animals was avoided (Macphail, 1996) (but see Rivalan et al., 2011 work). In the absence of such information the final performance usually worsens; therefore, the exploration phase typically lengthens and the optimal strategy is hardly found in 100 trials (Balodis et al., 2006; Fernie and Tunney, 2006; Glicksohn et al., 2007). However, when more trials are allowed, many individuals performing poorly in the first 100 trials are able to achieve good final performance (Balodis et al., 2006; Buelow et al., 2013; Bull et al., 2015). To that purpose, the number of trials was increased from 100 to 200. A full description of the electronic version of the IGT is given in Giustiniani et al. (2015).

2.5. Mouse gambling task

DM was evaluated using an mGT adapted from published protocols (Pittaras et al., 2016; de Visser et al., 2011). The experiment took place in a completely opaque four-arm radial maze (identical and equidistant arms, 37 cm long and 5.7 cm wide), with the common central zone used as a start-point. Mice were rewarded with grain-based pellets (20 mg Dustless Precision Pellets® Grain-Based Diet, PHYMED s.a.r.l., Paris, France) or punished with grain-based pellets previously treated with quinine (180 mM quinine hydrochloride, Sigma-Aldrich, Schnelldorf, Germany). Quinine pellets were poorly palatable but edible.

Prior to every experimental session, the animals were acclimatised to the behavioural room for 30 min. The experimental design was composed of 5 blocks of 20 trials over five days (a total of 100 trials per animal). The first ten trials of each block took place during the morning, the second ten trials during the afternoon. Before the first trial of the first block, mice had 3 min to explore and eat inside the maze (*first habituation period*). From the second block, mice had 2 min to explore the maze before the first trial, but no food was available (*general habituation period*). These habituation periods aimed to reduce the stressing effect of the

animals' first exposure to the dispositive. After this habituation, mice were placed at the start-point inside an opaque cylindrical structure to avoid early orientation. The cylinder was removed after 5 s and animals were allowed to choose an arm. Mice had one minute to choose an arm, explore it and eat the reward. An extra-minute was given if the choice was not made in time.

Our mGT was adapted in order to minimise the effect of satiety during the task. For that, two arms gave access to a small reward (1 pellet) in the first trial of each half block (trials 1, 11, 21, 31, 41, 51, 61, 71, 81 and 91) and bigger rewards (3-4 pellets) in the other 18 choices of each block with a small probability of presenting a punishment (3-4 quinine pellets, twice in 18 possible choices) (*advantageous arms C' and D'*). Arm C' and D' provided a maximal amount of 50 and 66 rewards per block, respectively.

The other two arms (*disadvantageous arms A' and B'*) gave access to a bigger reward (2 pellets) during the first trials of each half block, but bigger punishments (4-5 quinine pellets) and a smaller reward-possibility (4-5 pellets, once in 18 possible choices) in the block's remaining 18 choices. Arm A' gave access to a maximal amount of 10 rewards per block, arm B' until 14. Between consecutive trials, the animals were placed in their home cages for 90 s. The localisation of advantageous and disadvantageous arms was randomised with different award and punishment sequences for each animal.

2.6. Determination of interindividual differences

A k-mean clustering method, already used in mouse gambling tasks (Pittaras, 2016; Pittaras et al., 2016), was applied to identify interindividual differences in both, human and mice populations. This enabled the automatic identification of the objects' optimum partition into a specific number of clusters, minimising and maximising the intra and inter-cluster variance, respectively (Timmerman et al., 2013). In accordance with the literature (Pittaras et al., 2016), the mean percentage of advantageous choices was calculated for the tasks' final 30% when performance was highly stable ($W: p < 0.01$ in mice and humans). The individual performances were then divided into three groups: *good*, *intermediate* and *poor decision-makers (DMs)*.

2.7. Choice behaviours: rigidity, flexibility, lose-shift and win-stay scores

Maximisation of benefits and the reduction of costs, which characterise optimal performances, require flexibly adapting contingencies in order to favourably orient future choices. To compare different DM strategies, behavioural measures of cognitive processes at the beginning and the end of the task have been calculated. We measured our populations' rigidity score by calculating the highest percentage of choice of a deck or arm. The flexibility score is then the proportion of switches from one deck or arm to another. The lose-shift score, as a measure of negative outcome aversion, was assessed by calculating the proportion of switches after a penalty outcome. Win-stay scores likewise reflect the choice of the same option after receiving a reward. Each behavioural measure was calculated at the beginning (first 40% of the task) and at the end of the experiment (last 40% of the task), according to previous studies (Pittaras et al., 2016).

2.8. Reward sensitivity

In humans, the Behavioural Inhibition and Activation System scales (BIS/BAS) allowed us to approach behavioural motivation (Rizvi et al., 2016). In mice, reward sensitivity was evaluated at

the end of the mGT using the sucrose preference test (adapted from Lutz et al., 2013). For details, see Supplementary data.

3. Data analysis

3.1. Whole group analyses

A population's overall performance in terms of the advantageous choices for the gambling tasks was analysed with Bayesian and frequentist approaches. To consider learning rates, the relative risk (the ratio of the 'risk' of success in both populations (humans/mice) - RR) was measured. The resulting proportions were compared by Beta laws, whose parameters are interpreted in terms of the number of successes and failures, establishing the probability law of success rates' ratio. A credibility interval, inside which a parameter has a given probability (CI), was then calculated.

Hereafter, in order to address the behavioural kinetics the performance was divided into five blocks, each representing 20% of the task. The performance of each block was compared to the chance level using Student's tests (*t*-tests). The evolution of the performance was assessed by the analysis of variance (ANOVA), with the factor being 20%-blocks. Differences between populations were analysed using a partially repeated ANOVA with species as the between-subject variable and 20%-blocks as the within-subject variable.

The evolution of choice behaviours scores was first compared for each group using a *t*-test, followed by a comparison of the overall population's scores by two-way repeated-measures ANOVA with factors species and time course (first and last 40% of the task).

BIS/BAS data in humans and sucrose preference in mice were analysed using the Kruskal-Wallis (KW) test (see Supplementary data). Correlations between the endpoint performance (percentage of advantageous choices in the last 30% of the task) and choice behaviours were also carried out. Comparisons were Bonferroni corrected when necessary to account for the multiple comparisons.

3.2. Interindividual data analysis

For each group, the evolution of the gambling performance was assessed by repeated-measures ANOVAs, with the factors being 20%-Blocks and clusters (*good*, *intermediate* and *poor DMs*), followed by Mann Whitney U (MW) tests to show subgroup differences two by two. Each interindividual distribution was also compared between mice and humans using an MW test.

We used Wilcoxon (W) tests to compare, for each subgroup, the evolution of the gambling performance relative to the chance level. The subgroup's proportions of each population were compared using the Chi-square test. Differences in the choice behaviour scores between the subgroups in mice and humans were assessed by KW tests for the beginning and the end of the task. These measures were also compared between the species using MW tests. All MW and W tests were Bonferroni corrected.

The statistical significance threshold of all the tests was set at $p < 0.05$.

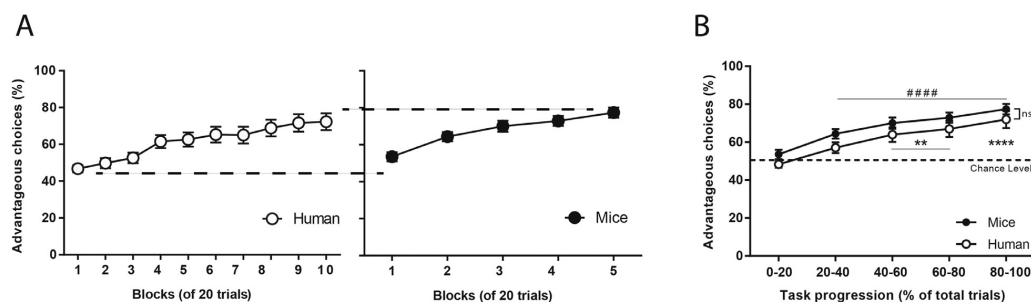


Fig. 1 Gambling performance in human and mice populations.

(A) Performances expressed as percentage of advantageous choices (mean \pm SEM) during task progression (blocks of 20 trials). After 100 trials, mice performed better than humans did (CI 95% = 1.191 to 1.277), but final performances were similar (CI 95% = 0.956 to 1.016). (B) No significant (ns) differences between species when comparing performance in 20% trial-blocks (repeated measures ANOVA). Human performance differed from chance level from the 3rd 20%-block onwards, while mice performance was already different from the 2nd block (t -test: mice: ###, $p < 0.0001$; human: **, $p < 0.01$; ****, $p < 0.0001$).

4. Results

4.1. Overall gambling performance: mice learn faster than humans (Fig. 1 and Supplementary Analysis 1)

The Bayesian analysis revealed a difference in the learning rate between the species (Fig. 1(A)), showing that mice learned faster than humans did: after 100 trials, mice had better overall gambling performance (CI 95% = 1.191 to 1.277). However, comparing the last 100 mice and human trials shows similar performances (CI 95% = 0.956 to 1.016). In order to circumvent variations in the gambling tasks' designs and to allow for the direct comparison of data, we analysed performance as a function of task progression, expressed as percentages. Splitting all trials into five 20% trial-blocks (Fig. 1(B)) clearly shows that humans acquired the task contingencies after completion of 40% of the task, performing above the chance level from the third 20%-block onwards (t -test, $p < 0.01$). A learning effect has been evidenced as the performance gradually improved over time ($F(4,156) = 15.4$; $p < 0.0001$) after 20% of the task was completed and till the end ($p < 0.0001$, above the behavioural output of the 1st 20%-block).

Mice, however, performed above the chance level already after the first 20% of the task, thus, orientating towards the favourable choices earlier than humans (t -test, $p < 0.0001$). The performance progressively improved over time ($F(4,195) = 11.9$; $p < 0.0001$) from the second 20%-block (2nd block: $p < 0.05$, 3rd block: $p < 0.001$, 4th and 5th blocks: $p < 0.0001$) and compared to the beginning of the task.

The comparison of the IGT and mGT performances shows no significant effect of the species in the percentage of advantageous choices ($F(1,78) = 3.1$; $p = 0.80$). Both populations performed similarly along the task progression, as shown by the absence of a statistical difference in the interaction between factors (species and 20%-Blocks) ($F(4,312) = 0.06$; $p = 0.99$).

These results show that both species learn the task contingencies, however demonstrating that mice are faster

learners than humans are. When compared as function of task progression, both populations display a similar improvement in performance and reach equivalent final performance despite the design differences.

4.2. Comparable categories of good, intermediate and poor DMs with similar proportions among the species (Fig. 2 and Supplementary Fig. 1 and Analyses 2 and 3)

In Figs. 2, A1 and B1 we plot individual data showing larger dispersion in humans than in mice. Intraspecies variability was investigated using k-mean clustering stratification according to the endpoint performance, discriminating three subpopulations: good, intermediate and poor DMs. Individuals majoratively displaying the optimal strategy were referred to as good DMs and represent 42.5% of the human population (mean percentage of advantageous choices \pm SEM: 97.4 ± 0.8) and 40% of the mice population (91.3 ± 1.5). Poor DMs remained at around 50% of advantageous choices, hence showing no significant preference for advantageous nor for disadvantageous options, represent 25% of humans (34.3 ± 5.2) and 22.5% of mice (54.1 ± 2.5). The third subpopulation corresponds to individuals that developed a preference towards some options -although they did not find the optimal strategy. These intermediate DMs represent 32.5% of humans (64.5 ± 2.4) and 37.5% of mice (73.8 ± 1.0). The proportions of each subpopulation in humans and mice were compared, and no significant differences were found (Chi-square test = 0.011, $p = 0.99$; confirmed by Bayesian analysis, see Supplementary Analysis 2) (Figs. 2, A2 and B2).

In humans, good DMs performed above the chance level after 40% of the task was completed (W , $p < 0.01$), while intermediate and poor DMs differed from the chance level only in the last 20%-block ($p < 0.05$). Furthermore, the ANOVA revealed a significant interaction between the clusters and 20%-blocks ($F(8,148) = 16.2$; $p < 0.0001$). Good DMs performed differently than poor DMs from the second

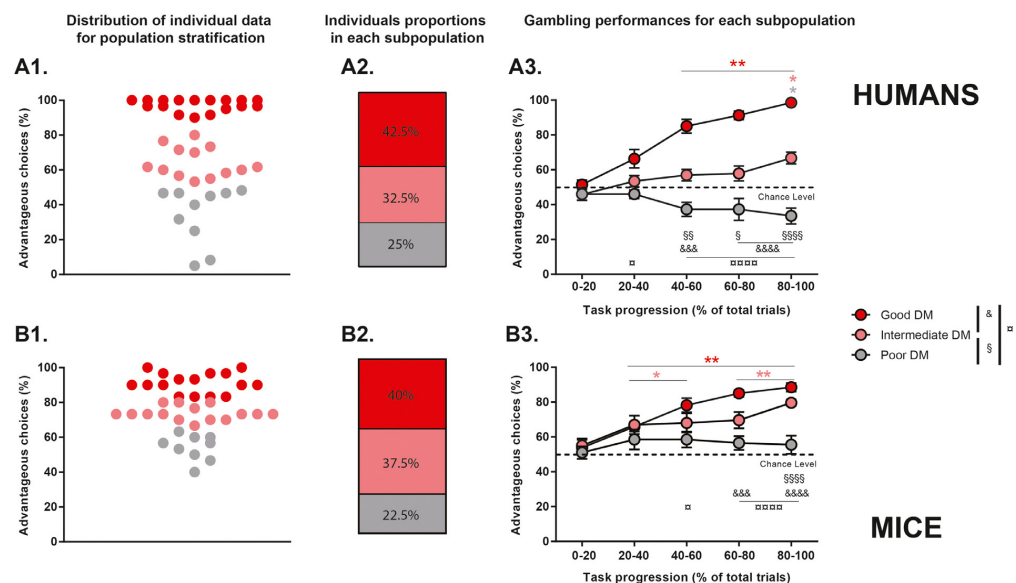


Fig. 2 Comparable categories of good, intermediate and poor DMs with similar proportions among species. Distribution in humans (A1) and mice (B1) of individual performance at the end of the task used for k-mean clustering. Proportions in humans (A2) and mice (B2) of individuals in good (red), intermediate (pink) and poor (grey) DM subpopulations (χ^2 , $p = 0.99$). Gambling performance in human (A3) and mice (B3) during task progression (20% trial-blocks). *W* tests to show group performance different from chance level (50%) (*, $p < 0.05$; **, $p < 0.01$). *MW* tests to show group differences (good vs intermediate: &, $p < 0.05$; &&, $p < 0.01$; &&&, $p < 0.001$; &&&&, $p < 0.0001$; good vs poor: □; intermediate vs poor: §). (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

20%-block onwards (*MW*, 2nd block: $p < 0.05$, 3rd to 5th blocks: $p < 0.0001$), and differently than intermediate DMs in the last 60% of the task (3rd block: $p < 0.001$, 4th and 5th blocks: $p < 0.0001$). Intermediate and poor DMs performed differently from the third block onwards (3rd block: $p < 0.01$, 4th block: $p < 0.05$, 5th block: $p < 0.0001$) (Figs. 2 and A3).

Regarding mouse data, intermediate and good DMs performed above the chance level soon after 20% of the task was completed (*W*, good 2nd to 5th blocks: $p < 0.01$; intermediate 2nd and 3rd blocks: $p < 0.05$, 4th and 5th blocks: $p < 0.01$), while poor DMs did not differ from the chance level ($p > 0.05$). The ANOVA also revealed a significant interaction between the factors ($F(8,185) = 5.3$; $p < 0.0001$). Mice from the good DM category performed differently than poor DMs after 40% of the task was completed (*MW*, 3rd block: $p < 0.05$, 4th and 5th blocks: $p < 0.0001$), and differently from the intermediate DMs later on (4th block: $p < 0.001$; 5th block: $p < 0.0001$). Mice from the intermediate and poor DM categories performed differently only during the last 20% of the task ($p < 0.0001$).

When comparing the gambling performance in mice and humans for each subpopulation, differences were found for intermediate DMs (*MW*, $p < 0.05$), with mice making more advantageous choices than humans. On the contrary, humans from the good DM category achieved more ad-

vantageous choices than mice of the same subgroup ($p < 0.01$). Furthermore, humans from the poor DM category made worse decisions than mice ($p < 0.01$). In addition, mice from the intermediate subgroup developed a weaker preference for the most advantageous option (D') than good DMs did, whilst humans from the same category chose the best deck (D) as good DMs did (see Supplementary Fig. 1 and Analysis 3).

Collectively these data reveal that both mice and human populations clustered into three comparable DM categories with closely matching proportions. However, the performance dispersion was more pronounced in the human population.

4.3. Relationship between the DM performance and choice behaviours in humans and mice (Fig. 3 and Supplementary Figs. 2 and 3)

4.3.1. Correlations between the endpoint performance and behavioural determinants of DM (Supplementary Fig. 2)

In humans, rigidity scores significantly correlated with the final performance at the beginning ($r = 0.428$, $p < 0.05$) and at the end of the task ($r = 0.489$, $p < 0.01$). Flexibility scores showed a negative correlation with the endpoint

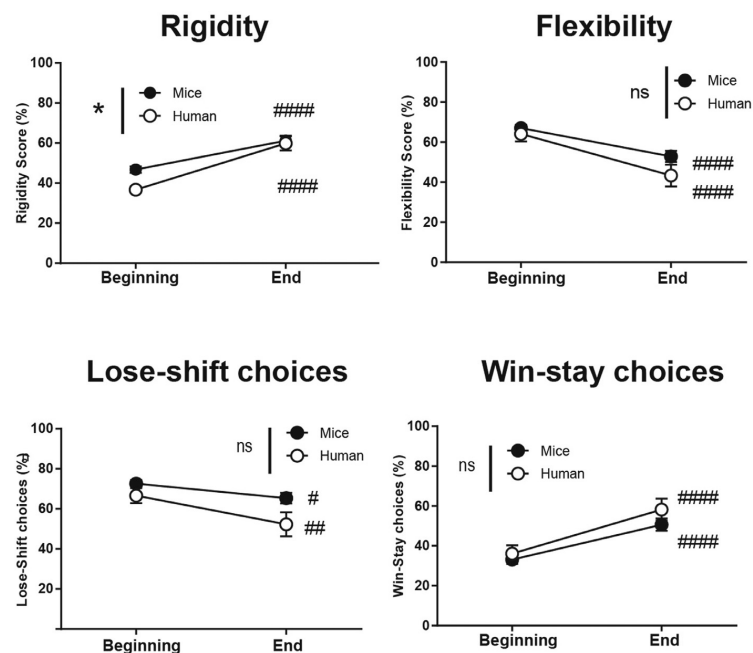


Fig. 3 Evolution of choice behaviors during task progression.

Significant progression of the rigidity, flexibility, lose-shift and win-stay choices for human and mice populations, between the beginning and the end of the task. Comparison of choice strategies at the population level by repeated measures ANOVA and post hoc *t*-tests (beginning vs end of the task: #, $p < 0.05$; ##, $p < 0.01$; ####, $p < 0.0001$; humans vs mice: *, $p < 0.05$).

performance at the beginning of the task ($r = -0.509$, $p < 0.01$), which disappeared at the end ($p > 0.05$). No significant correlation was found for the lose-shift choices and the endpoint performance, independently of the moment of the experiment. However, the win-stay choices at the beginning of the task significantly correlated with the endpoint performance ($r = 0.518$, $p < 0.01$).

In mice, like in humans, rigidity scores significantly correlated with the endpoint performance at the end of the task ($r = 0.669$, $p < 0.001$) but not at the beginning ($p > 0.05$). On the contrary, no correlation was found for flexibility and the endpoint performance at the beginning of the task ($p > 0.05$), but a negative correlation was found at the end ($r = -0.595$, $p < 0.001$). In the same way as humans, no correlation was found regarding the lose-shift choices ($p > 0.05$), but a significant correlation between the win-stay choices and endpoint performance at the end of the task ($r = 0.582$, $p < 0.001$).

In brief, DM strategies in humans and mice seem to rely on comparable adaptive choice behaviours that correlate with the endpoint performance.

4.3.2. Evolution of choice behaviours during task progression (Fig. 3 and Supplementary Fig. 3)

In both populations, a significant effect of the time course was found for flexibility ($F(1,78) = 48.4$, $p < 0.0001$), lose-shift ($F(1,78) = 15.8$; $p < 0.001$) and the win-stay scores ($F(1,78) = 51.9$, $p < 0.0001$) (Fig. 3).

Whereas no global difference among the species was found for any parameters, a significant interaction between the factors (species and time course) was found for rigidity ($F(1,78) = 4.6$, $p < 0.05$). Both populations became more rigid along the task (*t*-test, $p < 0.0001$), humans were significantly less rigid than mice at the beginning ($p < 0.05$) but not at the end of the experiment ($p = 1$) (Fig. 3).

The interspecies comparisons in all subpopulations revealed that, at the beginning of the task, humans from the intermediate and poor DM categories were significantly less rigid than the corresponding mice (MW, intermediate: $p < 0.0001$; poor: $p < 0.01$) but not good DMs ($p > 0.05$). At the end of the task, this property only remained preserved for humans from the intermediate DM subgroup ($p < 0.01$).

Flexibility scores were similar in both species for the intermediate and poor DMs at the beginning and the end of the experiment ($p > 0.05$). However, humans from the good DM category remained less flexible than mice of the same subgroup throughout the experiment (beginning: $p < 0.05$; end: $p < 0.001$). No subgroup differences in the lose-shift and win-stay choices were observed between the populations at the beginning of the task ($p > 0.05$). In the end, only mice from the good DM category were more prone to switch from an option after a penalty than humans ($p < 0.05$). However, humans from the good DM category continued choosing the same option after a positive outcome more frequently than mice ($p < 0.01$). No differences were found for the other subgroups ($p > 0.05$) (see Supplementary Fig. 3).

Overall, these data show that humans and mice do not differ in their choice strategies at the population level, but suggest there are slight differences in the intermediate and poor DMs subpopulations among the species. Remarkably, mice and humans from the good DM category behaved alike.

5. Discussion

The goal of this study was to directly compare DM under uncertainty between humans and mice using IGT adaptations according to the literature (van den Bos et al., 2006; Giustiniani et al., 2015; Pittaras et al., 2016). Our study shows closely related performance curves when comparing data as a function of task progression, suggesting that our animal model is a suitable candidate to bridge the findings between preclinical and clinical data. Therefore, our study offers a chance to get a better insight into the common conserved cognitive processes among mammals.

Humans rewarded with money and mice rewarded with food started the task with an explorative search, displaying an equal preference for either the advantageous or disadvantageous options. A preference for advantageous choices progressively emerged in both populations during what has been referred to as the exploitation phase (Daniel et al., 2017; Rivalan et al., 2013; de Visser et al., 2011). The overall performance was very similar, however, mice more promptly selected advantageous options than humans. The learning curves observed in our mice confirm previous animal results obtained in similar conditions (Pittaras, 2013; Pittaras et al., 2016) and in other variants of rodent gambling tasks (van Enkhuizen et al., 2014; de Visser et al., 2011).

Mice being faster than humans in developing a favourable strategy could be explained in terms of the rewards' nature, which is considered a major limitation for animal versions of gambling tasks. Modelling loss of reward in animals in a similar manner as in humans is a challenge. Food aids in the survival of species: it is a primary reinforcer since it strengthens behaviour and satiates the basic biological drives. Money, however, is a secondary reinforcer: its value is relative to the primary reinforcer. Hunger and satiety are factors difficult to control, which patently influence animals motivational state (van den Bos et al., 2014; Brevers et al., 2013). However, due to its subjective nature, the interest in money is also difficult to control, which raises similar concerns and can lead to similar consequences. This is the reason why the performance alignment in our mice and human participants is relevant.

In addition to the challenge associated with the reward nature and processing, the internal state (see the *somatic marker theory*, Verdejo-Garcia and Bechara, 2009) and context also generate differences in behaviour. Animal automated operant testing would also help deepen the analysis of motivational aspects for instance (Nithianantharajah et al., 2015).

A good validity of the model should also be substantiated by similar choice strategies, which can be investigated from additional cognitive proxies sub-serving behaviour (Winstanley and Clark, 2015). A closer look at the endpoint performance revealed that they correlated with the choice behaviours similarly but not identically in both popula-

tions. In fact, as advantageous choices increased along the task progression, preference for one option progressively emerged (increased rigidity), while fewer options were explored (decreased flexibility) with individuals becoming more sensitive to the reward (increased win-stay choices) and tending to more easily cope with penalties (decreased lose-shift choices).

Our results also highlighted close common interindividual variability in mice and humans when clustered into three subgroups of individuals, ranging from good over intermediate to poor DMs. Whether this clustering reflects the speed of acquiring a favourable strategy, or actually to different behavioural strategies could be further investigated (by a computational modelling analysis, for instance). However, the similar matched proportions of these subgroups for both species strengthen the face validity of our animal model.

Good DMs composed the largest subgroup in both species. Several human studies correlate a good performance predominantly with the development of an optimal strategy, accepting risky options (Barbalat et al., 2010; Charpentier et al., 2017). In mice studies a good performance is related to a secure strategy (Pittaras et al., 2016). Interestingly, mice from this category needed less time than humans to perform above chance level. Regarding the other subgroups, humans from the good DM category developed a stronger preference for one option over the intermediate and poor DMs, whereas mice of the same category differed only from the poor DMs. The evolution of the penalty aversion was also similar for both species, suggesting that the cognitive strategies underlying the DM performance might be similar in both species, at least for the good DM subgroup.

Individuals conforming to the intermediate DM category, while selecting advantageous choices more often, maintained a high level of exploration of all the options but did not achieve the best strategy to maximise their rewards. However, IGT studies have shown that the performance can be significantly enhanced with additional trials (Overman and Pierce, 2013). Furthermore, mice from this category became more sensitive to a reward along the task progression, whereas humans did not. This suggests that intermediate categories might not completely overlap between the species.

The subgroup with the worst performance, poor DMs, maintained the exploration of all the available options, exhibiting high behavioural flexibility. Nevertheless, they did not manage to find a favourable strategy along the task, suggesting an ineffective exploration of options. However, humans from this category ended the task performing significantly below the chance level in terms of the advantageous choices, which did not happen in the animal population. This kind of deleterious preference has also been described in mice following singled-session mGT protocols (van Enkhuizen et al., 2014) and other rodent IGT versions (Rivalan et al., 2009, 2013). Mice performing poorly in the gambling tasks have been proposed as models for the vulnerability of pathological gambling or addiction (van den Bos et al., 2013; Pittaras et al., 2016), in the same line of clinical studies seeking behavioural markers of pathological predisposition or endophenotypes (Cavedini et al., 2010; Zhang et al., 2015). Indeed, these animals seem less risk-averse, a trait that has been already

interpreted as an indicator of weaker cognitive control over immediate loss (de Visser et al., 2011).

Concerning the endpoint performance, interindividual variabilities show a larger spread in the human population, accounting for better and worse decisions. The extreme upper scores in humans might be a consequence of a longer gambling design (Balodis et al., 2006; Buelow et al., 2013; Bull et al., 2015). Food restriction in our animals could also explain why they never preferred the disadvantageous options. Noticeably, very low performance more reminiscent of human data than ours (Rivalan et al., 2009, 2013), illustrates that performance output highly depends on task design.

A general common pattern between humans and mice subpopulations has not been fully revealed when the endpoint performance is correlated with choice behaviours. By themselves, rigidity, flexibility and sensitivity to positive and negative outcomes cannot explain the evolution of the performance, neither their emergence, in the three subgroups equally. The IGT alone, unfortunately, does not allow to distinguish reward maximisation from ambiguity aversion for instance; its output is insufficient to determine why a participant selected an option. In this perspective, further behavioural characterisation has been attempted. In humans, we evaluated the motivation to avoid aversive and to approach goal-oriented outcomes, respectively by the behavioural inhibition and activation system's scales (BIS/BAS). The endpoint performance did not correlate with the BIS/BAS scores in humans (data not shown). In parallel, the reward sensitivity assessed by the sucrose preference task in mice did not significantly differ between the subgroups (see Supplementary Fig. 4). These results contrast with those from Pittaras et al. (2016), who described a stronger sucrose preference for good decision-makers ("safe") compared to mice which perform poorly ("risky"). Moreover, reward sensitivity did not correlate with the endpoint performance either (data not shown), whereas poor performance was likely mediated by sensitivity to a high reward in a single session mGT (van Enkhuizen et al., 2014). These apparent discrepancies could be accounted for by protocol variations and suggest complex relationships between reward sensitivity and DM strategies (van Enkhuizen et al., 2014).

A similar overall performance and comparable gambling strategies observed between both species suggest that they share conserved cognitive processes essential for successful DM. A wide range of comparative approaches, in rodents and in humans, have been proposed to discern between these processes during cognitive tasks, both at the attentional and the mnemonic level (Steckler and Muir, 1996). The literature has also shown that there is a strikingly similar range of cognitive abilities between rodents and humans, as well as a remarkably high degree of anatomical overlap in their brain functions (Woolley et al., 2013). Rodents are even able to outperform humans in some learning tasks (Vermaercke et al., 2014), but we do believe that our slightly differing kinetics are mainly accounted for by the task design variability (including the rewards' nature, the number of trials and the tasks 'probabilistic schedule).

To conclude, our data thus far suggest that in DM, mice behave in a way similar to humans: they tend to choose the option with the best long-term payoff more often as

the task progresses. Our results point to close patterns of choice behaviours present across species, but the parameters we evaluated are insufficient to draw a firm conclusion for the relationship between reward maximisation and risk aversion. Extensive behavioural characterisations and validated animal models will be crucial to study brain regions and circuits involved in DM, and the relationships between reward and cognitive control systems. Nonetheless, our results directly support good face validity of the mouse version of IGT. Future studies including pathological animal models, manipulations with pharmacological tools or brain stimulation techniques and computational modelling should help disentangle processes sub-serving choice strategies, clarifying DM in animals and humans.

Role of the funding source

This study was supported by grants from the Communauté d'Agglomération du Grand Besançon (LC), the French Eastern Interregional Group of Clinical Research and Innovation (GIRCI Est; appel à projet « jeunes chercheurs », APJ 2015; JG), which had no role in the study design; collection; analysis or interpretation of the data; writing of the report and in the decision to submit the paper for publication.

Contributors

LC, JG, DF, DG, YP conceived and designed the study. LC, JG, TC, BR, CJ, DG, MN, LP, EH performed the experiments. LC, DG, YP wrote the draft of the manuscript. All authors contributed to and have approved the final manuscript.

Conflict of interest

All authors declare no conflicts of interest.

Authors' disclosures

None.

Acknowledgements

The authors thank Dr Ruud van den Bos and Dr Sylvie Granon for their valuable advice in setting up the mGT. We are also grateful to Mr Hervé Reyssie from the *Animal Facilities, Besançon*, for his technical support, and Chrystelle Vidal for statistical support with Bayesian analysis.

Supplementary materials

Supplementary material associated with this article can be found, in the online version, at doi:10.1016/j.euroneuro.2019.11.005.

References

- Balodis, I.M., MacDonald, T.K., Olmstead, M.C., 2006. Instructional cues modify performance on the Iowa Gambling Task. *Brain Cognit.* 60, 109-117.
- Barbalat, G., Domenech, P., Vernet, M., Fournier, P., 2010. [Risk-taking in adolescence: a neuroeconomics approach]. *L'Encephale* 36, 147-154.
- Barry, D., Petry, N.M., 2008. Predictors of decision-making on the Iowa Gambling Task: independent effects of lifetime history of substance use disorders and performance on the Trail Making Test. *Brain Cognit.* 66, 243-252.
- Bechara, A., 1997. Deciding advantageously before knowing the advantageous strategy. *Science* 275, 1293-1295.
- Bechara, A., Damasio, A.R., 2005. The somatic marker hypothesis: a neural theory of economic decision. *Games Econ. Behav.* 52, 336-372.
- Bechara, A., Damasio, H., 2002. Decision-making and addiction (part I): impaired activation of somatic states in substance dependent individuals when pondering decisions with negative future consequences. *Neuropsychologia* 40, 1675-1689.
- Bechara, A., Damasio, A.R., Damasio, H., Anderson, S.W., 1994. Insensitivity to future consequences following damage to human prefrontal cortex. *Cognition* 50, 7-15.
- Bechara, A., Dolan, S., Denburg, N., Hindes, A., Anderson, S.W., Nathan, P.E., 2001. Decision-making deficits, linked to a dysfunctional ventromedial prefrontal cortex, revealed in alcohol and stimulant abusers. *Neuropsychologia* 39, 376-389.
- Bechara, A., Dolan, S., Hindes, A., 2002. Decision-making and addiction (part II): myopia for the future or hypersensitivity to reward? *Neuropsychologia* 40, 1690-1705.
- Bechara, A., Tranel, D., Damasio, H., 2000. Characterization of the decision-making deficit of patients with ventromedial prefrontal cortex lesions. *Brain* 123 (Pt 11), 2189-2202.
- van den Bos, R., Koot, S., de Visser, L., 2014. A rodent version of the Iowa Gambling Task: 7 years of progress. *Front. Psychol.* 5, 203. doi:10.3389/fpsyg.2014.00203.
- van den Bos, R., Lasthuis, W., den Heijer, E., van der Harst, J., Spruijt, B., 2006. Toward a rodent model of the Iowa gambling task. *Behav. Res. Methods* 38, 470-478.
- van den Bos, R., Davies, W., Deltu-Hagedorn, F., Goudriaan, A.E., Granon, S., Homberg, J., Rivalan, M., Swendsen, J., Adriani, W., 2013. Cross-species approaches to pathological gambling: a review targeting sex differences, adolescent vulnerability and ecological validity of research tools. *Neurosci. Biobehav. Rev.* 37, 2454-2471.
- Brand, M., Recknor, E.C., Grabenhorst, F., Bechara, A., 2007. Decisions under ambiguity and decisions under risk: correlations with executive functions and comparisons of two different gambling tasks with implicit and explicit rules. *J. Clin. Exp. Neuropsychol.* 29, 86-99.
- Buelow, M.T., Okdie, B.M., Blaine, A.L., 2013. Seeing the forest through the trees: improving decision making on the Iowa gambling task by shifting focus from short- to long-term outcomes. *Front. Psychol.* 4, 1-10. doi:10.3389/fpsyg.2013.00773.
- Buelow, M.T., Suhr, J.A., 2009. Construct validity of the Iowa Gambling Task. *Neuropsychol. Rev.* 19, 102-114.
- Brevers, D., Bechara, A., Cleeremans, A., Noël, X., 2013. Iowa Gambling Task (IGT): twenty years after - gambling disorder and IGT. *Front. Psychol.* 4, 665. doi:10.3389/fpsyg.2013.00665.
- Bull, P.N., Tippett, L.J., Addis, D.R., 2015. Decision making in healthy participants on the Iowa Gambling Task: new insights from an operant approach. *Front. Psychol.* 6, 1-17. doi:10.3389/fpsyg.2015.00391.
- Cavedini, P., Zorzi, C., Piccinni, M., Cavallini, M.C., Bellodi, L., 2010. Executive dysfunctions in obsessive-compulsive patients and unaffected relatives: searching for a new intermediate phenotype. *Biol. Psychiatry* 67, 1178-1184.
- Charpentier, C.J., Aylward, J., Roiser, J.P., Robinson, O.J., 2017. Enhanced risk aversion, but not loss aversion, in unmedicated pathological anxiety. *Biol. Psychiatry* 81, 1014-1022.
- Daniel, M.L., Cocker, P.J., Lacoste, J., Mar, A.C., Houeto, J.L., Belin-Rauscent, A., Belin, D., 2017. The anterior insula bidirectionally modulates cost-benefit decision-making on a rodent gambling task. *Eur. J. Neurosci.* 46, 2620-2628.
- Davis, C., Fox, J., Patte, K., Curtis, C., Strimas, R., Reid, C., McCool, C., 2008. Education level moderates learning on two versions of the Iowa Gambling Task. *J. Int. Neuropsychol. Soc.* 14, 1063-1068.
- van Enkhuizen, J., Geyer, M.A., Young, J.W., 2013. Differential effects of dopamine transporter inhibitors in the rodent Iowa gambling task: relevance to mania. *Psychopharmacology* 225, 661-674.
- van Enkhuizen, J., Henry, B.L., Minassian, A., Perry, W., Milienne-Petiot, M., Higa, K.K., Geyer, M.A., Young, J.W., 2014. Reduced dopamine transporter functioning induces high-reward risk-preference consistent with bipolar disorder. *Neuropsychopharmacology* 39, 3112-3122.
- de Visser, L., Homberg, J., Mitsogiannis, M., Zeeb, F.D., Rivalan, M., Fitoussi, A., Galhardo, V., van den Bos, R., Winstanley, C.A., Deltu-Hagedorn, F., 2011. Rodent versions of the Iowa gambling task: opportunities and challenges for the understanding of decision-making. *Front. Neurosci.* 5, 1-21. doi:10.3389/fnins.2011.00109.
- Fernie, G., Tunney, R.J., 2006. Some decks are better than others: the effect of reinforcer type and task instructions on learning in the Iowa gambling task. *Brain Cognit.* 60, 94-102.
- Festing, M.F.W., 2014. Evidence should trump intuition by preferring inbred strains to outbred stocks in preclinical research. *ILAR J.* 55, 399-404.
- de Froment, A.J., Rubenstein, D.I., Levin, S.A., 2014. An extra dimension to decision-making in animals: the three-way trade-off between speed, effort per-Unit-Time and accuracy. *PLoS Comput. Biol.* 10, e1003937.
- Giustiniani, J., Gabriel, D., Nicolier, M., Monnin, J., Haffen, E., 2015. Neural correlates of successful and unsuccessful strategic mechanisms involved in uncertain decision-making. *PLoS One* 10, e0130871.
- Glicksohn, J., Naor-Ziv, R., Leshem, R., 2007. Impulsive decision-making: learning to gamble wisely? *Cognition* 105, 195-205. doi:10.1016/j.cognition.2006.08.003.
- Heilbronner, S.R., 2017. Modeling risky decision-making in nonhuman animals: shared core features. *Curr. Opin. Behav. Sci.* 16, 23-29.
- Lehner, R., Balsters, J.H., Herger, A., Hare, T.A., Wenderoth, N., 2017. Monetary, food, and social rewards induce similar Pavlovian-to-instrumental transfer effects. *Front. Behav. Neurosci.* 10, 1-12. doi:10.3389/fnbeh.2016.00247.
- Lutz, P.E., Reiss, D., Ouagazzal, A.M., Kieffer, B.L., 2013. A history of chronic morphine exposure during adolescence increases despair-like behaviour and strain-dependently promotes sociability in abstinent adult mice. *Behav. Brain Res.* 243, 44-52.
- Macphail, E.M., 1996. Cognitive function in mammals: the evolutionary perspective. *Brain Res. Cognit. Brain Res.* 3, 279-290.
- Meyer, G., 2004. Neuroendocrine response to casino gambling in problem gamblers. *Psychoneuroendocrinology* 29, 1272-1280.
- Meyer, G., Hauffa, B.P., Schedlowski, M., Pawlak, C., Stadler, M.A., Exton, M.S., 2000. Casino gambling increases heart rate and salivary cortisol in regular gamblers. *Biol. Psychiatry* 48, 948-953.
- Miedl, S.F., Fehr, T., Meyer, G., Herrmann, M., 2010. Neurobiological correlates of problem gambling in a quasi-realistic blackjack scenario as revealed by fMRI. *Psychiatry Res. Neuroimaging* 181, 165-173.

- Nithianantharajah, J., McKechnie, A.G., Stewart, T.J., Johnstone, M., Blackwood, D.H., St Clair, D., Grant, S.G.N., Bussey, T.J., Saksida, L.M., 2015. Bridging the translational divide: identical cognitive touchscreen testing in mice and humans carrying mutations in a disease-relevant homologous gene. *Sci. Rep.* 5, 14613.
- Overman, W.H., Pierce, A., 2013. Iowa gambling task with non-clinical participants: effects of using real + virtual cards and additional trials. *Front. Psychol.* 4, 1-15. doi:10.3389/fpsyg.2013.00935.
- Pittaras, E., 2013. Mice gamble for food: individual differences in risky choices and prefrontal cortex serotonin. *J. Addict. Res. Ther.* 1 (S4:011), 1-9. doi:10.4172/2155-6105.S4-011.
- Pittaras, E., 2016. Marqueurs Comportementaux Et Neurochimiques Individuels De La Prise De Décision Chez La Souris Et Effets D'une Dette De Sommeil. Université Paris-Saclay.
- Pittaras, E., Callebert, J., Chennaoui, M., Rabat, A., Granon, S., 2016. Individual behavioral and neurochemical markers of unadapted decision-making processes in healthy inbred mice. *Brain Struct. Funct.* 221, 4615-4629.
- Prendergast, B.J., Onishi, K.G., Zucker, I., 2014. Female mice liberated for inclusion in neuroscience and biomedical research. *Neurosci. Biobehav. Rev.* 40, 1-5.
- Rivalan, M., Ahmed, S.H., Dellu-Hagedorn, F., 2009. Risk-prone individuals prefer the wrong options on a rat version of the Iowa Gambling Task. *Biol. Psychiatry* 66, 743-749.
- Rivalan, M., Coutureau, E., Fitoussi, A., Dellu-Hagedorn, F., 2011. Inter-Individual decision-making differences in the effects of cingulate, orbitofrontal, and prelimbic cortex lesions in a rat gambling task. *Front. Behav. Neurosci.* 5, 1-10. doi:10.3389/fnbeh.2011.00022.
- Rivalan, M., Valton, V., Seriès, P., Marchand, A.R., Dellu-Hagedorn, F., 2013. Elucidating poor decision-making in a rat gambling task. *PLoS One* 8, e82052.
- Rizvi, S.J., Pizzagalli, D.A., Sproule, B.A., Kennedy, S.H., 2016. Assessing anhedonia in depression: potentials and pitfalls. *Neurosci. Biobehav. Rev.* 65, 21-35.
- Singh, V., 2016. Sex-Differences, handedness, and lateralization in the Iowa Gambling Task. *Front. Psychol.* 7, 1-15. doi:10.3389/fpsyg.2016.00708.
- Steckler, T., Muir, J.L., 1996. Measurement of cognitive function: relating rodent performance with human minds. *Brain Res. Cognit. Brain Res.* 3, 299-308.
- Steingrover, H., Wetzels, R., Horstmann, A., Neumann, J., Wagenmakers, E.-J., 2013. Performance of healthy participants on the Iowa Gambling Task. *Psychol. Assess.* 25, 180-193.
- Suhr, J.A., Tsanadis, J., 2007. Affect and personality correlates of the Iowa Gambling Task. *Personal. Individ. Differ.* 43, 27-36.
- Timmerman, M.E., Ceulemans, E., De Roover, K., Van Leeuwen, K., 2013. Subspace K-means clustering. *Behav. Res. Methods* 45, 1011-1023.
- Tuttle, A.H., Philip, V.M., Chesler, E.J., Mogil, J.S., 2018. Comparing phenotypic variation between inbred and outbred mice. *Nat. Methods* 15, 994-996.
- Verdejo-García, A., Bechara, A., 2009. A somatic marker theory of addiction. *Neuropharmacology* 56, 48-62.
- Vermaercke, B., Cop, E., Willems, S., D'Hooge, R., Op de Beeck, H.P., 2014. More complex brains are not always better: rats outperform humans in implicit category-based generalization by implementing a similarity-based strategy. *Psychon. Bull. Rev.* 21, 1080-1086.
- de Visser, L., Baars, A.M., Lavrijsen, M., van der Weerd, C.M.M., van den Bos, R., 2011. Decision-making performance is related to levels of anxiety and differential recruitment of frontostriatal areas in male rats. *Neuroscience* 184, 97-106.
- Winstanley, C.A., Clark, L., 2015. Translational models of gambling-related decision-making. In: Robbins, T.W., Sahakian, B.J. (Eds.), *Translational Neuropsychopharmacology*. Springer International Publishing, Cham, pp. 93-120.
- Woolley, D.G., Laeremans, A., Gantois, I., Mantini, D., Vermaercke, B., Op de Beeck, H.P., Swinnen, S.P., Wenderoth, N., Arckens, L., D'Hooge, R., 2013. Homologous involvement of striatum and prefrontal cortex in rodent and human water maze learning. *Proc. Natl. Acad. Sci.* 110, 3131-3136.
- Young, J.W., van Enkhuizen, J., Winstanley, C.A., Geyer, M.A., 2011. Increased risk-taking behavior in dopamine transporter knockdown mice: further support for a mouse model of mania. *J. Psychopharmacol.* 25, 934-943.
- Zeeb, F.D., Winstanley, C.A., 2013. Functional disconnection of the orbitofrontal cortex and basolateral amygdala impairs acquisition of a rat gambling task and disrupts animals' ability to alter decision-making behavior after reinforcer devaluation. *J. Neurosci.* 33, 6434-6443.
- Zeeb, F.D., Robbins, T.W., Winstanley, C.A., 2009. Serotonergic and dopaminergic modulation of gambling behavior as assessed using a novel rat gambling task. *Neuropsychopharmacology* 34, 2329-2343.
- Zhang, L., Dong, Y., Ji, Y., Zhu, C., Yu, F., Ma, H., Chen, X., Wang, K., 2015. Dissociation of decision making under ambiguity and decision making under risk: a neurocognitive endophenotype candidate for obsessive-compulsive disorder. *Prog. Neuropsychopharmacol. Biol. Psychiatry* 57, 60-68.

