

**Neuroengineering of hallucinations through fMRI
neurofeedback:
Understanding and modulating neural dynamics for
the control of hallucinations in health and disease**

Présentée le 21 avril 2023

Faculté des sciences de la vie
Chaire Fondation Bertarelli de neuroprothétique cognitive
Programme doctoral en neurosciences

pour l'obtention du grade de Docteur ès Sciences

par

Herberto DHANIS PEDRO DE BARROS CAMACHO

Acceptée sur proposition du jury

Prof. K. Hess Bellwald, présidente du jury
Prof. O. Blanke, Prof. D. N. A. Van De Ville, directeurs de thèse
Prof. J. Sulzer, rapporteur
Prof. G. Allali, rapporteur
Prof. V. Zerbi, rapporteur

to my family, who has taught me so many things...

Abstract

Recent years have seen considerable progress in understanding brain function, and in modulating it through real-time fMRI neurofeedback (NF). Inspired by these advances, paired with the unique possibility of inducing the clinically relevant *Presence Hallucination* (PH: strange sensation of having someone behind when no one is there) through MR-compatible robotics, I set out to: (1) understand the brain dynamics that lead to PH and have participants regulate them as to modulate PH; (2) study the neural underpinnings of this type of hallucinations in Parkinson's disease (PD); and (3) advise on the optimal setup to induce PH.

In Part I of my thesis (Study 1), I investigated the neural correlates of PH-induction, hypothesizing that underlying hallucinations were sporadic dysfunctions in temporal processes of brain activity. Using Co-Activation Pattern (CAP) analysis I identified patterns of activity and studied their occurrence and transition probabilities, while participants experienced robot-induced PH (riPH). With this I showed that sensitivity to riPH depended on a temporary shift in transition probabilities that caused all CAPs to increase transitions to a specific brain pattern (PH-network).

With this knowledge, I then paired the MR-compatible PH-induction system with fMRI-NF (Part I – Study 2), to provide informative feedback on the PH-network activity as PH was induced, and allow participants to achieve volitional control over it. During three NF-training days, participants learned to up-regulate and down-regulate the PH-network, which lead to an increase in sensitivity to riPH post-training, as compared to pre-training. Moreover, for participants that were successful during NF and became sensitive to riPH, we noted lasting changes in brain activity marked by an increased occurrence of the PH-network during induction.

In Part II (Study 3), I investigated the neural correlates of PH in PD, as it is a common hallucination in this condition that might predict cognitive decline and persistent psychosis. I investigated riPH, cognitive impairment and fMRI neural correlates, in a cohort of patients stratified based on the severity of hallucinations: no hallucinations, minor hallucinations (subgroup including PH), and structured hallucinations (mostly visual). I showed increased sensitivity to riPH across the axes of hallucination severity and cognitive impairment. Studying multivariate patterns of brain activity and behavior, I then identified that antagonistic activations and deactivations between large brain networks important for cognition, self-related processing, and vision, underpinned preserved cognitive capabilities, decreased sensitivity to riPH, and no hallucinations.

Finally in Part III (Study 4), motivated by the translation of riPH from healthy to clinical populations, I analyzed all experiments that used riPH, and quantified the effect of various experimental parameters

Abstract

with a Bayesian meta-analysis. With this I was able to propose setup recommendations for specific purposes.

At the junction of hallucination theories, induction, temporal dynamics of brain processing, and brain regulation through fMRI-NF, my thesis identified neural underpinnings of hallucinations in health and disease, and further showed these can be subject of volitional regulation. Together these findings advance the causal understanding of brain dynamics in neuropsychiatry and could have major translational applications for novel anti-psychotic fMRI therapies.

Keywords: Hallucinations, presence hallucination, fMRI, real-time fMRI, neurofeedback, Parkinson's disease, minor hallucinations, dynamic functional connectivity, Co-Activation Patterns (CAPs)

Resumé

Ces dernières années, des progrès considérables ont été réalisés dans la compréhension des fonctions cérébrales et, en retour, dans la modulation de cette fonction cérébrale par neurofeedback (NF) par IRMf en temps réel. Inspiré par ces progrès, ainsi que par la possibilité unique d'induire l'hallucination de présence (HP: sensation étrange d'avoir quelqu'un derrière soi alors qu'il n'y a personne) par le biais d'une robotique compatible avec l'IRM, j'ai entrepris de: (1) comprendre la dynamique cérébrale qui conduit à l'hallucination de présence et d'amener les participants à la réguler, afin de moduler la PH; (2) d'étudier les fondements neuronaux de ce type d'hallucinations dans la maladie de Parkinson (MP); et (3) optimiser la méthodologie pour obtenir une configuration idéale pour induire la HP.

Dans la première partie de ma thèse (Partie I – Étude 1), j'ai étudié les corrélats neuronaux de l'induction de la HP, en émettant l'hypothèse que les hallucinations sous-jacentes étaient des dysfonctionnements sporadiques dans les processus temporels de l'activité cérébrale. À l'aide de l'analyse des modèles de cartes de coactivation (CCA), j'ai identifié des modèles d'activité et étudié leur occurrence et les probabilités de transition, tandis que les participants faisaient l'expérience d'une HP induite par un robot (HP-r : HP robotique). J'ai ainsi montré que la sensibilité à la HP-r dépendait d'un changement temporaire des probabilités de transition qui faisait que tous les CCA augmentaient les transitions vers un modèle cérébral spécifique (réseau-HP).

Fort de ces connaissances, j'ai ensuite associé le système d'induction de HP compatible avec l'IRM à une IRMf-NF (Partie I - Étude 2), afin de fournir un retour d'information sur l'activité du réseau-HP au fur et à mesure de l'induction de HP, et de permettre aux participants d'exercer un contrôle volontaire sur celle-ci. Pendant trois jours d'entraînement au NF, les participants ont appris à réguler à la hausse et à la baisse le réseau PH. Cela a conduit à une augmentation de la sensibilité à la HP-r après la formation, par rapport à la formation préalable. De plus, pour les participants qui ont réussi la NF et sont devenus sensibles à la HP-r, nous avons noté des changements durables dans l'activité cérébrale marqués par une augmentation de l'occurrence du réseau HP pendant l'induction.

Dans la deuxième partie (Partie II – Étude 3), j'ai étudié les corrélats neuronaux de la HP dans la MP, car il s'agit d'une hallucination courante dans cette maladie qui pourrait prédire le déclin cognitif et la psychose persistante. J'ai étudié la HP-r, les troubles cognitifs et les corrélats neuronaux de l'IRMf, dans une cohorte de patients stratifiés en fonction de la gravité des hallucinations: pas d'hallucinations, hallucinations mineures (sous-groupe incluant la HP), et hallucinations structurées (principalement visuelles). J'ai montré une sensibilité accrue au HP-r sur les axes de la sévérité des hallucinations et des troubles cognitifs. En étudiant les schémas multivariés de l'activité cérébrale et du comportement, j'ai ensuite identifié que des activations et désactivations antagonistes entre les

Resumé

grands réseaux cérébraux importants pour la cognition, le traitement lié à l'individu et la vision, sous-tendaient des capacités cognitives préservées, une sensibilité réduite à la HP-r et l'absence d'hallucinations.

Enfin, dans la troisième partie (Partie III – Étude 4), motivé par la transposition de HP-r des populations saines aux populations cliniques, j'ai analysé toutes les expériences qui ont utilisé HP-r, et quantifié l'effet de divers paramètres expérimentaux avec une méta-analyse bayésienne. J'ai ainsi pu proposer des recommandations de configuration à des fins spécifiques.

À la jonction des théories de l'hallucination, de l'induction, de la dynamique temporelle du traitement cérébral et de la régulation cérébrale par IRMf-NF, ma thèse a identifié les fondements neuronaux des hallucinations dans la santé et la maladie, et a montré que celles-ci pouvaient être volontairement régulé. Ensemble, ces résultats font progresser la compréhension causale de la dynamique cérébrale en neuropsychiatrie et pourraient avoir des applications translationnelles majeures pour de nouvelles thérapies anti-psychotiques par IRMf.

Mots-clés: Hallucinations, hallucination de présence, imagerie par résonance magnétique fonctionnelle (IRMf), IRMf en temps réel, neurofeedback, maladie de Parkinson, hallucinations mineures, connectivité fonctionnelle dynamique, cartes de coactivation

Acknowledgments

First and foremost, I want to thank my thesis director, Professor Olaf Blanke, for being an incredible supervisor during my doctoral studies. By means of interesting and motivating discussions, he was a great source of inspiration. By trusting me with enormous independence to pursue my ideas, he helped develop the many skills I have and helped me become the scientist I am today. Ultimately, I thank him for believing in me and my work. I would further say that Olaf has taught me how to see *beyond*, have an eagle-eye perspective on research, and importantly, not just how to report my findings but how to *tell a story*. These are rare and very important skills. I feel very lucky, and am truly grateful, for these and many other lessons, under Olaf's supervision.

I must also thank Dimitri Van De Ville, the co-director of my thesis. The discussions with Dimitri opened the door to a world of methodologies that excited and kept my mind running about what we could explore and uncover with them. I am very grateful for his methodological supervision and for his passion for the field. Also, I am very grateful to both, for the kind understanding, when I was faced with unfortunate health-related circumstances.

A big thanks is also in order to Jevita Potheegadoo and Fosco Bernasconi. To Jevita, for the master organization skill that kept our clinical project running, but also for her help, advice, kindness, and nuanced guidance. To Fosco, for many interesting discussions, for his inspiringly competent look into data, being a great person to discuss in-depth neuroscience and methods, and for help and advice during my PhD.

Then, I would also like to thank Nathan and Giulio, for the guidance at the beginning of my PhD, Sara for the shared interest in dynamic connectivity methods that allowed for great discussions, Cyrille for the help with patients, Sophie Oblette and Alice Goodman for warm admin support, and more broadly to the people at the Blanke Lab for the amazing time I spent here, with a special mention to the "new" PhDs, Louis, Neza and Juliette for renewed motivation and great spirits. Furthermore, I also extend my thanks to Nico, Ellie, and Lukas for interesting discussions on fMRI neurofeedback, Thomas Bolton for the kind methodological help at the beginning of my thesis, and Loan and Roberto from the MRI facility with whom I spent countless hours.

Now, there is a very special group of people I absolutely must thank. The team that I started my PhD with, that made this time absolutely fantastic, some that I have the privilege to consider my best friends, Giedre, Nathalie, Julia, Pavo, Matteo, Giuliana, Eva, Myeong, and Hyukjun. I could not have

Acknowledgments

wished for a better integration in Switzerland. Through really so many experiences together, you turned this PhD into an amazing adventure.

Pavo, who's basically my LNCO brother (since we met actually during our master thesis before our PhDs), I must thank for countless moments of in-depth discussions about neuroscience and how we would change the field, to insightful personal conversations, and of course the best trips/parties/retreats, you name it and he's up for it in the best mood! I thank Julia for showing me the way into Swiss culture, making sure I learned how to ski, how to go up a hill in my ski boots (yes it was a hill!), and for having been always so extremely welcoming. Life in Switzerland would have been very different without her. The three of us have been through so much, and you have been by my side in a very special period of my life. I am truly thankful for the two of you (the forever weird triangle).

Then my partners in crime, Giedre and Nathalie! Nathalie, my eternal and fabulous dance partner, PhD sister, truly one of the best and kindest persons I have had the pleasure of meeting, I really thank you for your support in so many aspects of my life. Giedre, with whom it is so easy to share life, discuss its most deep aspects, be them good or bad, with whom it is so easy to just *be* with, I am so thankful that we have grown to know each other throughout this time! It's a unique sensation to "find your people" and the three of us have shared so many amazing moments together. I can't wait for everything we will explore and conquer in the future!

A very deep thank you also goes to one of my most cherished friends, karaoke and party partner, James King. Jim, you have been a guide through so many aspects of life I had never thought reachable, and you have been an absolute rock when health was not shining bright. I truly thank you from the bottom of my heart.

Por fim, gostaria de agradecer à minha família. Aos meus tios e padrinhos, Bela e Jorge, por terem sido presenças ativas na minha vida. À minha prima Carolina, que considero como irmã, pelo apoio incondicional em várias partes da minha vida, pelos bons momentos partilhados e experiências ao longo da vida. Ao meu pai, Kelso, por me ter ensinado valores importantes de bondade, compaixão para com os outros, amizade, pela sua imensa inteligência emocional, aceitação, e pelo seu apoio. À minha mãe Maria João, por me ter passado desde tenra idade o amor pela ciência (e estética), sem o qual certamente nunca teria seguido a via académica, bem como pelo apoio e muitos sacrifícios que fez por mim e pela minha educação. E finalmente à minha avó Maria Alice, cujas qualidades absolutamente impressionantes de serenidade, paz, bondade e inteligência são e serão sempre faróis de esperança na minha vida.

To finish up, I thank my family. My godparents, Bela and Jorge, for having been such active presences in my life. My cousin Carolina, whom I consider a sister, for her unconditional support in many aspect of my life, for the good moments shared, laughter, and so many experiences during our lives. To my father Kelso, for teaching me important values of kindness, compassion, friendship, for his emotional intelligence, acceptance and for his support. To my mother, Maria João, who from a young age passed on to me the love for science (and aesthetics), without which I certainly would have never pursued an academic career. I thank her as well for the support and many sacrifices done for me and my education. Finally, to my grandmother, Maria Alice, whose impressive personal qualities of serenity, peace, kindness, understanding, and intelligence, are and will always be a beacon of hope in my life.

Herberto Dhanis

25th January 2023

Contents

Abstract.....	iii
Resumé	v
Acknowledgments.....	vii
Contents.....	ix
List of Figures	xiii
List of Tables	xv
Motivational Preface.....	3
Introduction	5
1.1 Brain mapping and functional magnetic resonance imaging.....	5
1.1.1 Physical Principles of fMRI	7
1.1.2 Quantifying brain activity.....	8
1.2 Hallucinations	11
1.2.1 Deficits in self-monitoring: Cancellation theory	11
1.2.2 Presence Hallucination.....	14
1.2.3 Parkinson's Disease	17
1.3 Neuroimaging of Hallucinations	19
1.3.1 Evidence for dysconnectivity in hallucinations	19
1.3.2 Dynamic brain processes	20
1.3.3 Dynamic brain processes in Parkinson's Disease	21
1.3.4 Neuroimaging evidence related to the Presence Hallucination	22
1.4 Real-time fMRI neurofeedback	23
1.4.1 Neuropsychiatric applications of real-time fMRI neurofeedback.....	24
1.4.2 Methodological considerations in real-time fMRI neurofeedback.....	25
1.5 Thesis at a glance.....	26

Contents

1.6	Overview and personal contributions	27
Part I: Identifying and modulating neural correlates of PH		29
2.1	Study 1: Robotically-induced hallucination triggers subtle changes in network transitions	31
2.1.1	Abstract	33
2.1.2	Introduction	35
2.1.3	Materials and methods	36
2.1.4	Results	42
2.1.5	Discussion.....	48
2.1.6	References.....	53
2.1.7	Supplementary Information.....	59
2.2	Study 2: Sustained bidirectional self-regulation of hallucination networks through real-time fMRI neurofeedback	73
2.2.1	Abstract	75
2.2.2	Introduction	77
2.2.3	Results	79
2.2.4	Discussion.....	87
2.2.5	Methods	90
2.2.6	References.....	97
2.2.7	Extended Data	102
2.2.8	Supplemental Information	104
Part II: Neural correlates of hallucination severity in Parkinson's Disease		127
3.1	Study 3: Robot-induced hallucinations and PLS-CAPs reveal neural correlates of neurocognitive-psychiatric preservation in PD	129
3.1.1	Abstract	131
3.1.2	Introduction	133
3.1.3	Materials and methods	134
3.1.4	Results	138
3.1.5	Discussion.....	141
3.1.6	References.....	145
Part III: Quantifying and optimizing PH induction		151
4.1	Mega-Analysis of Presence Hallucination Induction Experiments Using Robotically Mediated Sensorimotor Conflicts	153
4.1.1	Abstract	155
4.1.2	Introduction	157

Contents

4.1.3 Results	159
4.1.4 Discussion.....	166
4.1.5 Methods	171
4.1.6 References.....	175
General Discussion.....	181
5.1 The importance of temporal dynamics	181
5.1.1 Indirect evidence for hierarchical dynamics: Meta-states.....	183
5.2 Control of hallucinations	184
5.3 Neuropsychiatry of Parkinson's Disease	185
5.4 Conclusion and outlook.....	186
References	189
Appendix	209
6.1 TbCAPs: A toolbox for co-activation pattern analysis	211
6.2 Thought Consciousness and source monitoring depend on robotically controlled sensorimotor conflicts and illusory states.....	237
6.3 Presence Hallucination during locomotor activities in patients with Parkinson's Disease	269
6.4 Neuroscience robotics for controlled induction and real-time assessment of hallucinations	275
Curriculum Vitae	311

Contents

List of Figures

Figure 1.1 Neuroimaging techniques and their relative invasiveness, temporal and spatial resolutions	6
Figure 1.2 Illustration of the physical properties of (de/oxygenated blood	8
Figure 1.3 Superior views of the CEN, SN, and DMN	9
Figure 1.4 Example of static versus dynamic functional connectivity analysis	10
Figure 1.5 Forward models for motor control and consequences in motor action performance.....	12
Figure 1.6 Example of Bayesian weighting of prior and sensory evidence.....	13
Figure 1.7 Robotic setup to induce PH and induction results	16
Figure 1.8 Progression of hallucinations and insight/cognition.....	19
Figure 1.9 Triple Network Model	21
Figure 1.10 Network interactions between the DMN, CEN and a visual network in hallucinations in PD	22
Figure 1.11 Real-time fMRI neurofeedback	24
Figure 2.1 Robotic system and experimental paradigm.....	37
Figure 2.2 Occupancy and average duration of the CAPs.....	43
Figure 2.3 Anatomy of CAP 6 (PH) and CAP 9 (PE).....	44
Figure 2.4 CAPs' transition dynamics across different conditions	45
Figure 2.5 PH and PE transitions dynamics	47
Figure 2.6 [S1] Consensus clustering	59
Figure 2.7 [S2] CAPs from 1(top) to 9 (bottom) after spatial z-scoring.....	60
Figure 2.8 [S3] CAP transition matrixes for the various conditions	65
Figure 2.9 [S4] CAP transitions in the asynchronous condition for participants who experienced induced PH and for those who did not	66
Figure 2.10 [S5] CAP transitions in the asynchronous condition for participants who experienced induced PE and for those who did not.....	67
Figure 2.11 [S6] Consensus clustering (control - PCC)	68
Figure 2.12 [S7] Consensus clustering (control - mSFG and aINS).....	69
Figure 2.13 Experimental Paradigm	78
Figure 2.14 Assessment of Neurofeedback Regulation Performance.....	80
Figure 2.15 High-regulation performance in neurofeedback and transfer runs ..	81

List of Figures

Figure 2.16 Anatomy of the PH-network and of other identified brain states	82
Figure 2.17 Re-instantiation mechanism during the up-regulation NF-condition of HRP runs	83
Figure 2.18 Avoidance mechanism during the control condition of HRP runs	84
Figure 2.19 PH ratings before and after NF training	85
Figure 2.20 Difference in PH state occurrences when comparing after vs. before training	86
Figure 2.21 [Extended Data 1] Other brain states identified through CAP analysis	102
Figure 2.22 [Extended Data 2] Occurrences for the posterior state during NF runs	102
Figure 2.23 [Extended Data 4] Loss of Agency ratings before and after NF training	103
Figure 2.24 [Extended Data 3] PE ratings before and after NF training	103
Figure 2.25 [S2] Consensus clustering for different number of centroids	104
Figure 2.26 [S1] Stability of consensus clustering for different number of centroids (k)	104
Figure 2.27 [S3] Change in transition probabilities when comparing HRP versus non-HRP runs during up-regulation	110
Figure 2.28 [S4] Change in transition probabilities when comparing the up-regulation condition between HRP and non-HRP transfer runs	112
Figure 2.29 [S5] Occurrences for the different states during transfer runs (HRP vs non-HRP runs)	114
Figure 2.30 [S6] Ratings for the control questions of the PH-questionnaire	114
Figure 3.1 Experimental paradigm	135
Figure 3.2 Results of the sensorimotor delay dependency task	138
Figure 3.3 Identified Co-Activation Patterns	139
Figure 3.4 PLS-CAPs analysis	140
Figure 4.1 Estimates for the effect of asynchrony across all experiments	160
Figure 4.2 Posterior estimates for the experimental parameters modulating riPH	161
Figure 4.3 Posterior estimate for age and trait characteristics of participants ..	162
Figure 4.4 Estimation of the effect of age and PDI across the ordinal scale	163
Figure 4.5 Characteristics of the ordinal rating scale used for riPH	164
Figure 4.6 Mediation analysis for the effect of sensitivity to different sensations	165
Figure 4.7 Posterior estimates for the sensitivity mediators	166
Figure 5.1 Example of meta-states	183

List of Tables

Table 2.1 Questionnaire on robot-induced sensations	59
Table 2.2 [S2] CAPs' cluster peaks	61
Table 2.3 [S3] Occupancy statistics	64
Table 2.4 [S4] Average duration statistics	64
Table 2.5 [S5] Hierarchical models for transition probabilities	65
Table 2.6 [S6] Average duration statistics (control - PCC)	68
Table 2.7 [S7] Average duration statistics (control - mSFG and aINS)	70
Table 2.8 [S8] Comparison between different models for transition probabilities (control - mSFG and aINS)	71
Table 2.9 [S1] Cluster peaks for the extracted CAPs	105
Table 2.10 [2] Results of the transition probability analysis for NF runs	109
Table 2.11 [S3] Intermediate results for the transition probability analysis of NF runs	109
Table 2.12 [S4] Final analysis for the transition probabilities of different brain states during up-regulation or rest for NF runs	110
Table 2.13[S5] Results for the unrestricted transition probability analysis during transfer runs	111
Table 2.14 [S6] Restricted analysis of transition probabilities during transfer runs	111
Table 2.15 [S7] Results of the analysis of occurrences of brain states during NF runs	112
Table 2.16 [S8] Intermediate results for the analysis of occurrences of brain states during NF runs	112
Table 2.17 [S9] Final results for the analysis of occurrences of brain states during NF runs	113

List of Tables

Table 2.18 [S10] Results for the unrestricted analysis of occurrences of brain states during transfer runs	113
Table 2.19 [S11] Results for the restricted analysis of the occurrences of brain states during transfer runs	113
Table 2.20 [S12] Unrestricted analysis of differential occurrences of PH-task sessions	115
Table 2.21 [S13] Intermediate model for differential occurrences of PH-task sessions	115
Table 2.22 [S14] Intermediate model 2 for differential occurrence of PH-induction sessions	116
Table 2.23 [S15] Final model assessing differential occurrences of PH-induction sessions	116
Table 2.24 [S16] Control analysis for differential occurrences of passivity experiences (PE) – unrestricted model	116
Table 2.25 [S17] Control analysis for differential occurrences of PE - restricted model	117
Table 2.26 [S18] Control analysis for differential occurrences using Loss of Agency (LoA)	117
Table 2.27 [S19] Control analysis (using LoA) for differential occurrences of PH-induction sessions - restricted model 1	117
Table 2.28 [S20] Control analysis (using LoA) for differential occurrences of PH-induction sessions - restricted model 2	118
Table 2.29 [S21] Analysis of differential transition probabilities of PH-induction days - unrestricted	118
Table 2.30 [S22] Final model of differential transition probabilities of PH-induction days.....	119
Table 2.31 [S23] Control analysis (PE) for differential transition probabilities - unrestricted.....	120
Table 2.32 [S24] Control analysis (PE) for differential transition probabilities - unrestricted.....	121
Table 2.33 [S25] Control analysis (LoA) for differential transition probabilities - unrestricted.....	122
Table 2.34 [S26] Control analysis (LoA) for differential transition probabilities - unrestricted.....	123
Table 2.35 [S27] PH questionnaire	124

**“Reality, or at least our perception of
it, is a controlled hallucination”**

Anil Seth

Motivational Preface

Hallucinations represent a fascinating feat of the brain's intricate functioning – the creation of something, out of nothing. The brain adds to its own reality without any external stimulation. For the *hallucinator*, these creations of the brain can seem as real as this thesis you are reading. Why do they feel so real, if at the end of the day they are *just* conjurations of *a brain*? The truth is, the brain's interpretations of our senses are the only reality we know. In the words of Oliver Sacks: “one does not see with the eyes; one sees with the brain”. If that's the case, then what are the mechanisms leading up to these astonishing creations that are hallucinations? They could be abnormalities in localized brain areas, dysfunctional connections between parts of the brain, or perhaps more widespread changes. Regardless of what they might be, I would like you to ask yourself, could we learn to modulate those *erroneous* brain mechanisms and achieve control over hallucinations? Would such a thing be remotely possible... and if so, would there be any implications for the *hallucinator*?

With my thesis, I aimed to answer these three essential questions, focused around a specific hallucination that is the Presence Hallucination – the strong and convincing sensation of having someone behind you when no one is actually there (Brugger et al., 1996). Our interest with this particular hallucination fell on three important aspects. It is a hallucination that can occur in healthy individuals (Peter Brugger et al., 1999) and different clinical conditions (Fenelon et al., 2011; Llorca et al., 2016; Nagahama et al., 2007). It is a characteristic hallucination of Parkinson's disease that might predict worse functional outcomes (Factor et al., 2003; ffytche et al., 2017; Lenka et al., 2019). And perhaps most importantly for our goals, it is a hallucination that we can experimentally induce in the lab and by consequence study directly (Bernasconi et al., 2022; Blanke et al., 2014). In practice, this meant I could identify its brain mechanisms in healthy individuals while the hallucination was being induced, create an experiment to have participants control these mechanisms (and potentially the presence hallucination), and begin the necessary work to translate my findings to Parkinson's disease.

The challenge ahead was not easy. There was considerable evidence that hallucinations and psychosis in general, entailed a fundamental dysconnectivity between frontal and temporo-parietal regions of the brain (i.e. Stephan, et al., 2009), and the field was perhaps too focused on this. However, this seemed quite unspecific to *just* hallucinations and rather a trait of psychosis in general. Simultaneously, the field also started to recognize the importance of wider brain processes involving intrinsic brain networks (connections over many regions) in hallucinations (Menon, 2011). Arguably, a hallucination which typically occurs during limited time windows, could not *only* be explained by constant dysfunction, be it at a region or widespread level in the brain. I was rather interested in temporal processes of brain activity and how these might engage in some form of *temporary*

Motivational Preface

dysfunction, that results in a hallucination. This idea and the advent of various methods to probe fine temporal processes (dynamic functional connectivity) greatly helped in accomplishing my first objective: the identification of the neural correlates of presence hallucination.

The control of the brain mechanism behind this hallucination was done later, through a technique called neurofeedback. In brief, neurofeedback relies on the concept that for the brain to learn *anything*, it requires informative feedback (Sitaram et al., 2017). Take it playing the piano, one must hear when they press the wrong key, to know a mistake was made, and correct it. Even in this scenario, all one is doing is learning to control the brain activity that translates into the motor movements necessary to play the piano. In many other situations, including hallucinations, there is no signal which is *as overt* as the one in the example. Thus, neurofeedback uses techniques to measure brain activity in real-time, and provides a direct feedback to the user about their current mental state. Concretely, I measured brain activity in real-time using functional magnetic resonance imaging and relayed to my participants a wave-like sound, which conveyed information about how close (or not) their brain state was, to the one identified in my previous work and that corresponded to that of the presence hallucination. With this, my participants achieved volitional modulation of these brain processes and subsequent control of the hallucination

Finally, one of my motivation to develop this work was its potential relevance for hallucination therapies in Parkinson's disease, given that the presence hallucination is a characteristic hallucination in this condition (Fenelon et al., 2011). It is in fact part of a specific subset of minor hallucinations (ffytche et al., 2017; Goetz et al., 2006; Pagonabarraga et al., 2014), hypothesized to precede more complex ones, such as visual hallucinations (Lenka et al., 2019), which are in turn linked to cognitive decline in this patient population (Bejr-kasem et al., 2021; Morgante et al., 2012). While this is acknowledged in clinical practice, there is no standardized way of treating minor hallucinations. This is largely because patient's insight is retained for these, and the use of anti-psychotic medication in Parkinson is not trivial (ffytche et al., 2017). Consequentially, hallucinations tend to be addressed only when there is a pathological progression that is damaging to the patient. In fact, literature recommends "personal coping strategies" as a first-resort for hallucinations (Diederich et al., 2003). Hence, potentially stopping or delaying the progression of such minor hallucinations *could eventually* slow down cognitive decline or even the progression of psychiatric symptoms. So in itself, understanding the root neural correlates of these hallucinations was already a considerable endeavor, but importantly, in its broader scope that supersedes this thesis, laid the foundations for a neurofeedback therapy attempting to decrease the incidence of minor hallucinations, and delay the progression of neuropsychiatric symptoms in Parkinson's Disease.

Hopefully, by expressing the motivation that lead me to develop this PhD thesis, I have enticed your curiosity for this work, that I have produced alongside the two labs that I am affiliated with. Now, for you to delve in it, the next section will present you with the necessary background to follow along.

Introduction

In this section I will introduce the necessary topics to understand how this thesis was developed. I will start by introducing functional magnetic resonance imaging. This technique is central to this thesis as it allowed the search for temporal processes of hallucinations. Furthermore, the methods of analysis targeting them will also be discussed here. I start by giving a brief historical and methodological context to this technique, mostly because while nowadays we – especially in academic circles – talk about measuring brain activity as if it was trivial, we shouldn't forget that functional magnetic resonance imaging only has 30 years of history, and linking behavior to brain functioning was very different in the past. Following this, I will present relevant background in hallucinations, how they are important in Parkinson's disease, and then, of course, the presence hallucination, which is central to these studies. The following chapter will be dedicated to neuroimaging of hallucinations. In particular, I will focus on the theories surrounding the role of fundamental brain networks and their temporal processes in hallucinations, also with an emphasis on Parkinson's disease, and neuroimaging of presence hallucination. Finally, I will conclude with real-time functional magnetic resonance imaging neurofeedback, given that this thesis includes a major application of this method to the topics previously discussed, and constitutes a major motivator for our work in Parkinson's disease.

1.1 Brain mapping and functional magnetic resonance imaging

For a long time, a considerable amount of our knowledge about brain function and its links to behavior, was limited to two areas: animal studies and neurological/neuropsychological work. Through the former, it was possible to perform invasive studies and directly measure how behavior and exposure to stimuli (perception) modulated brain activity. Such studies were naturally indirect with regards to the functioning of the human brain, but some aspects were believed to be similar across humans and others animals. An example is the influential work of Hubel and Wiesel during the 1960s, which identified specific neurons in the brain's of cats, that react differently, to arrays of different patterns being displayed to the cats (for more details, see for example: Wurtz, 2009). While this avenue of research allowed for an extreme control of experimental conditions and brain activity (nowadays the reader can even refer to optogenetics), it also implied searching for fundamental mechanisms shared across species, if that work was to be applied to humans. In that sense, linking human behavior directly to brain function was a more complicated endeavor. In the past, this was many times only possible through the lens of rare and unfortunate events in someone's life (typically accidents or disease which would lead to changes in behavior). Two famous examples in the area of neurology, are the works of Pierre Paul Broca (circa 1861) and Carl Wernicke (circa 1881). Broca, accompanied two patients with almost complete loss of speech, and upon their death, identified in their

Introduction

autopsies, a brain lesion over the lateral frontal lobe. This led Broca to link the function of that brain region to speech production (see: Dronkers et al., 2007). Carl Wernicke through extensive work on medical records and autopsies of stroke patients identified the brain area responsible for language comprehension (see: Krestel, 2013). In both cases, the link between function and behavior was revealed posthumously and relied on focalized lesions and specific symptoms allowing a precise correlation. But the link between brain and behavior was not always necessarily achieved after death. Some rare cases studied within the field of neuropsychology allowed for these discoveries to be done while the patient was alive. One is the case of Phineas Gage (1848), a railroad construction worker whose left pre-frontal cortex was obliterated in an accident with a metal bar leaving him with significant personality changes (see: Damasio et al., 1994). Another is that of Henry Molaison (1953), widely known as patient H.M., in which an experimental neurosurgery attempting to stop epileptic seizures saw his medial temporal lobes removed, leaving him without the ability to form new memories (see: Squire, 2009).

It was only with the considerable technological development of the twentieth century, that the advent of non-invasive brain recordings occurred. With it, it finally became possible to record brain activity and link it to behavior in a manner that has revolutionized our understanding of its inner working. Nowadays, a diverse set of techniques makes use of distinct physical properties of the brain in order to record its activity. Electroencephalography (EEG) and its implanted version – electrocorticography (ECoG), detect the brain's electrical activity, while magnetoencephalography (MEG) measures small magnetic field changes stemming from the aforementioned electrical activity (Lopes da Silva, 2013). Functional near infrared spectroscopy (fNIRS; Boas et al., 2004) and functional magnetic resonance imaging (fMRI; Ogawa et al., 1990; Ogawa & Lee, 1990) on the other hand, measure relative changes in local concentrations of oxygen in the blood, while positron emission tomography (PET) focuses on measuring brain cells metabolism (Lameka et al., 2016). Crucially, all these techniques (amongst others) come with important advantages and disadvantages in terms of spatial resolution, temporal resolution, and invasiveness (Figure 1.1). For example, while EEG offers excellent

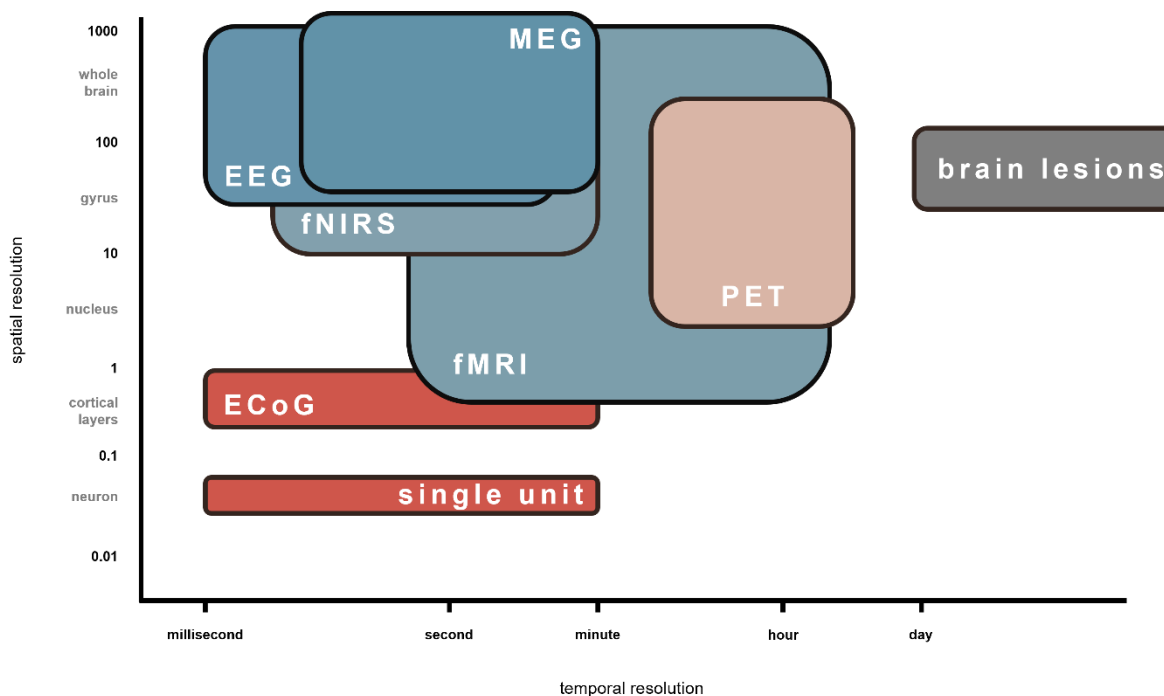


Figure 1.1 Neuroimaging techniques and their relative invasiveness, temporal and spatial resolutions

Temporal and spatial resolutions are shown for different imaging techniques. Techniques in red are invasive, while those in blue are non-invasive. PET is in between as it does involve the injection of a radioactive tracer. Brain lesions are shown for historical references. This figure is based on work by: Sejnowski and colleagues (2014), and Martin & Huettel (2013).

Introduction

temporal resolution, it lacks spatial resolution. ECoG has a much better spatial resolution, but requires surgical implantation.

Functional magnetic resonance imaging (fMRI), the technique central to this thesis, uses intrinsic properties of water molecules and neurovascular coupling – the relationship between neuronal activity raising metabolic needs, and an increase in blood supply to the area – to sample brain activity. Its use as a valid technique to acquire brain activity in humans, was first demonstrated in 1992 for both visual activity (Kwong et al., 1992) and sensorimotor activity (Bandettini et al., 1992). Depending on technical characteristics, a typical fMRI can resolve activity in the range of 1 to 3 millimeters, with a temporal resolution of around 700 msec to 2.5 seconds. Note that despite the temporal resolution that can be achieved nowadays, the fMRI signal has an intrinsic temporal limitation, termed haemodynamic delay, which is the time it takes for neuronal activity to trigger an increase in blood supply to the area (see next chapter).

1.1.1 Physical Principles of fMRI

In this section, I will give a brief introduction on the principles of acquisitions through magnetic resonance imaging and how they can be used to measure brain function. However, as it is not the goal of this thesis to detail on the fundamental physics of MRI a more curious reader might find more thorough information on this topic in a review by Heeger and Ress (2002), and in the references contained within.

Magnetic resonance imaging (MRI) makes use of intrinsic properties of the brain's composition, just like the techniques described above. Here, the properties explored are that of hydrogen atoms (protons) which are abundant in every tissue of our body as they are present in water molecules. The brain itself is approximately 80% water (Oros-Peusquens et al., 2019). To acquire an image through MRI, the MR-scanner imposes a static magnetic field which causes protons to align their spin with the direction of the magnetic field (van Geuns et al., 1999). A precise electromagnetic radiofrequency pulse is then used to excite the hydrogen atoms, causing them to precess around the direction of the magnetic field, and become perpendicular to it. Upon the end of this excitation phase, the pulse stops and the proton which is no longer being excited, precesses back to its original alignment with the magnetic field, emitting back the received energy (relaxation phase). Because different tissues will have different densities of protons, the obtained signal will be more intense in areas with higher density and vice-versa. By varying the time at which the MR-scanner measures the emission produced during the relaxation phase, it is then possible to contrast different tissues *.

To measure brain activity, fMRI focuses on a specific contrast called, Blood-Oxygen Level Dependent (BOLD) contrast, that is very sensitive to changes in the magnetic field. This is interesting, because particles with different properties will introduce different inhomogeneities in the magnetic field. For example, oxyhaemoglobin (red blood cells transporting oxygen) has diamagnetic properties, meaning it is more compliant with the magnetic field, and generates a stronger signal during MRI acquisition. Conversely, deoxyhaemoglobin (red blood cells without oxygen) has paramagnetic properties, meaning it is less compliant with the magnetic field, and hence generates a weaker signal (Figure 1.2). What ends up giving away the information on brain activity is a phenomenon called neurovascular coupling (Heeger & Ress, 2002), which as was briefly alluded to before, pairs neuronal metabolism and blood requirements. The idea is that as local brain activity increases so do the metabolic needs of those neurons, leading to an increase in the need for oxygen and nutrients, that boosts supply of arterial blood to that local area. This in turn is picked up

* Other important aspects exist such as the weighting of the images. T1-weighted images are sensitive to the time it takes for the proton to re-align with the magnetic field. T2-weighted images rather measure the time it takes for the alignment of the protons with the magnetic field to decay. While fMRI relies on a particular T2-weighting, these concepts are outside the scope of this thesis.

Introduction

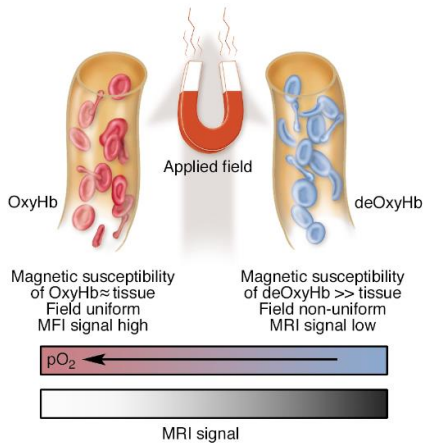


Figure 1.2 Illustration of the physical properties of (de/oxygenated blood
Given its diamagnetic properties, oxygenated blood does not interfere as much with the magnetic field as deoxygenated blood which is paramagnetic and hence introduces inhomogeneities in the magnetic field. This produces a higher signal in MRI measurements which is then used to inform about surrounding brain activity. Reproduced with permission from Gore (2003).

by the MRI-scanner precisely due to the changes in ratio of oxyhaemoglobin and deoxyhaemoglobin. Note however, that these changes are not instantaneous. While neuronal activity happens very fast (1-10 milliseconds range), the changes in blood supply to the activated area that are dependent on vasodilation (seconds range). Moreover, there is a washout period from metabolism generating more deoxyhaemoglobin to the newly supplied oxyhaemoglobin. This implies a haemodynamic delay between neuronal activity and the peak of the fMRI signal due to the flow of oxyhemoglobin, of about 4 to 6 seconds, which is not circumventable, as it is an inherent property of neurovascular coupling (Logothetis et al., 2001).

1.1.2 Quantifying brain activity

It is important now to understand how researchers use this to study brain activity. Most fMRI experiments are under the umbrella of two types of paradigms: task-based fMRI, in which participants perform a task during the acquisition, and resting-state fMRI, in which participants are mind-wandering in the MR-scanner for the short duration of 5 to 10 minutes (Gore, 2003). Both paradigms have been used in this thesis and I illustrate below some approaches to analyze such data and the questions they answer.

In a simple task, where participants perform finger-tapping in blocks of tasks (e.g. 30 seconds finger-tapping, 15 seconds of rest, repeated over time), activation-based analyses can identify the regions responsible for those movements. For this, the 3D fMRI data is processed by a statistical model, typically a general linear model (GLM) which identifies the voxels that are more active in the task periods (Friston et al., 1994). For a more complex task, such as the recognition of faces versus houses, one might opt to use multi-variate pattern analysis (MVPA) to identify patterns of brain activation that correspond to each of the stimuli (Haxby et al., 2001, 2011; Norman et al., 2006). Albeit different complexities, both these methods answer the question of *what* is more or less active during the task.

1.1.2.1 Static functional connectivity and large-scale brain networks

Another type of analysis that holds an important place in fMRI analysis is functional connectivity (FC). With this we explore if two or more brain regions of interest (ROI) have a correlated pattern of activity across the entire time of the acquisition, for example by computing the Pearson correlation between both ROIs timeseries of activity (K. J. Friston, 2011). When no a-priori hypothesis exists regarding connectivity, one might also pick a brain region to serve as a *seed-region* from which then functional connectivity is computed to the whole brain. The focus, now, is not only on the activity but on the *functional relationship* between different brain regions during task or during resting-state. A major

Introduction

contribution of functional connectivity analysis has been the insights provided with regards to the functional organization of the brain.

Biswal and colleagues (1995) were the first to use functional connectivity to identify a large-scale brain network at rest – in this case, a somato-motor network. Since then different studies (Damoiseaux et al., 2006; Fox et al., 2005; Fox & Raichle, 2007; Shirer et al., 2012) identified that the brain divides functional activity in groups of highly cooperating brain regions, which are referred to as large-scale brain networks (or resting-state networks, RSNs, or intrinsic brain networks; Yeo et al., 2011). Of particular interest for this thesis are three fundamental large-scale brain networks with central roles in function (Figure 1.3). The default-mode network (DMN), believed to be central to aspects of mind-wandering and self-related processing (Beckmann et al., 2005; Buckner et al., 2008; Raichle et al., 2001); the central executive network (CEN), responsible for sustained attention, executive control, and cognitive manipulation tasks (Fox et al., 2006; Seeley et al., 2007); and the Salience Network (SN; also known as ventral attention network) responsible for reacting to sensitive

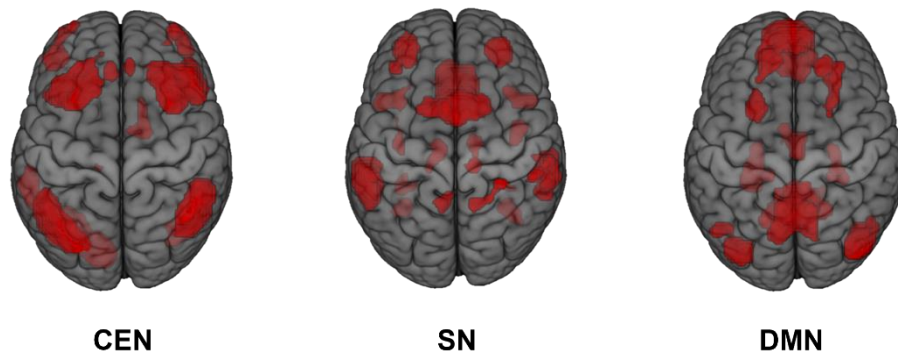


Figure 1.3 Superior views of the CEN, SN, and DMN

Superior views for the CEN, SN, and DMN are shown based on data from Shirer, et al. (2012)

stimuli and balancing the activity of other large-scale brain networks (Ham et al., 2013; Seeley et al., 2007; Sridharan et al., 2008).

1.1.2.2 Dynamic Functional Connectivity

The methods described above lead to significant progress in our understanding of brain activity, however, they had the intrinsic limitation of treating FC as static, that is, assuming their relational properties did not change over time. However, FC does fluctuate over time (Chang & Glover, 2010), including for the large-scale brain networks identified above (Chen et al., 2016). Methods that try to identify these changes in FC over time are under the umbrella of dynamic functional connectivity (dFC; Preti et al., 2017; see Figure 1.4). A notable approach to dFC is called sliding window analysis and is performed by breaking down the time-course into smaller windows and measuring functional connectivity inside these smaller time windows, as opposed to the entire acquisition time (Hutchison et al., 2013; Kucyi & Davis, 2014). One of its key assumptions is that connectivity fluctuates slowly over time. However, others have put forward that tapping into dFC properties can be done by focusing on briefs moments of high brain activity (Tagliazucchi et al., 2011). This method has shown to be able to recover patterns of known brain networks and raised interesting considerations about brain functioning. Namely the potential implication that the brain spends a considerable amount of time close to a critical point, that when reached creates an avalanche of activity in which relevant information is condensed (Tagliazucchi et al., 2012).

1.1.2.3 Co-Activation Pattern Analysis and temporal modeling

Co-Activation Pattern (CAP) analysis is a two-step procedure based first on the point process proposed by Tagliazucchi and colleagues (2012), followed by a clustering procedure using k-means (Liu & Duyn, 2013). First, one or more relevant seed regions are chosen, and their time courses are analysed to identify timepoints in which their activity exceeds a certain threshold. Then, whole-brain data from those timepoints are selected and clustered using k-means to identify a K number of CAPs. The optimal K can be determined using a stability metric called consensus clustering (Monti et al., 2003). Following this, timepoints assigned to the same cluster are averaged to obtain representative CAPs. In this manner, a set of CAPs represents individual groups of brain regions that exhibit synchronous activity at different points in time.

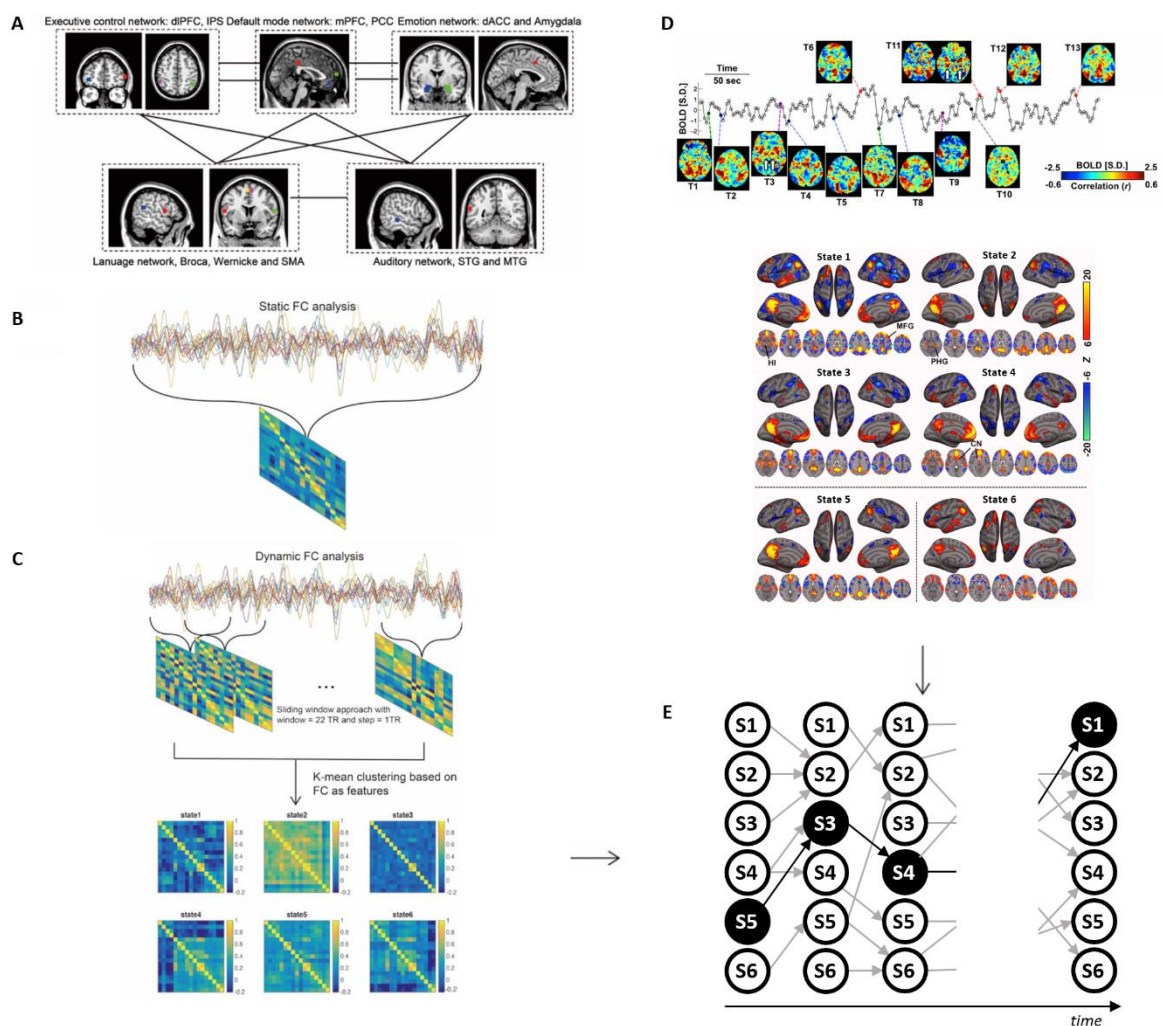


Figure 1.4 Example of static versus dynamic functional connectivity analysis

(A) Example of hypothetical brain activity involving the CEN, DMN, SN, and areas of a language network, and an auditory network. (B) Activity recovered through static functional connectivity analysis. A Pearson correlation is computed between all voxels, taking into account the entire period of brain activity. Only one correlation matrix is recovered (C) The same activity recovered using dynamic connectivity through the sliding-window method. Multiple correlation matrixes are recovered and k-means is used to find centroids across the identified windows leading to 6 dynamic connectivity matrixes. (D) Activity recovered using dynamic connectivity methods, but this time with CAPs analysis. Activity of a seed region is used to select time-point, which will have specific patterns at each time-point as seen for T1, T2, and so on. The patterns of each time-point are clustered using k-means to recover a number of centroids which are representative CAPs. (E) For both methods (C) and (D) temporal modelling of the matrixes/CAPs can be performed. Adapted with permission from Geng, et al. 2020, and Liu, et al. 2013, under the Creative Commons License 4.0.

Introduction

In this procedure we can note two interesting aspects that are advantageous. First, CAPs makes use of the assumptions of the point-process, but by applying them to a seed region, can focalize the search to periods associated to expected functions of the seed region. Second, by focalizing on selecting time points based on the activity of a seed, it allows to detect periods of de-activation, which was not possible using the point-process at the level of the whole-brain.

While using CAPs already disentangles some of the dynamic structure of FC (for example by identifying sub-states of the large-scale brain networks Liu & Duyn, 2013; for extended review see Liu et al., 2018), in this thesis, I have further explored their temporal relationships, by analysing their occurrences over conditions of interest, and by studying their relationship to each other through the analysis of transition probabilities from one CAP to any other (Chen et al., 2015).

With this, the main focus of the analyses of this thesis has been on identifying the large intrinsic brain networks at play in our tasks and patients, as well as their temporal relationships, which will be detailed in the next chapters.

1.2 Hallucinations

Hallucinations are defined as perceptual experiences occurring without any corresponding external stimuli (Allen et al., 2008). For example, hearing a voice without any sound being present (auditory hallucinations; Tracy & Shergill, 2013), or seeing something without anything actually being there (visual hallucinations; Manford & Andermann, 1998), amongst others. Importantly they are different from illusions, which are a misinterpretation of a stimuli that is actually there (Eagleman, 2001).

The spectrum in which hallucinations occur is quite large. They can arise in healthy individuals, over the course of their lifetime, with a prevalence of around 5 to 10% (Linscott & van Os, 2010; Majer et al., 2018; Nuevo et al., 2012). In this population, anxiety is known to slightly increase the prevalence of hallucinations (Larøi et al., 2019), and bereavement in particular can lead to a significant prevalence of hallucinations of presence (as much as 80%, Hayes & Leudar, 2016).

Outside healthy populations, hallucinations are fairly relevant as clinical markers of disease onset and progression. They are common in psychiatric conditions, such as schizophrenia with a prevalence of up to 80% of auditory hallucinations and up to 50% of visual hallucinations (Millan et al., 2016), and can mark the clinical onset of the disease (Menezes et al., 2006); and are also common in neurodegenerative diseases such as dementia with Lewy bodies (DLB) and Parkinson's Disease (PD), two conditions marked by significant prevalence of visual hallucinations of 80% (McKeith et al., 2017) and 70% (Diederich et al., 2009; Ffytche et al., 2017), respectively. Neurological incidents leading to ischemia (strokes, infarctions) can also temporarily lead to hallucinations if the sensory cortexes are affected (De Haan et al., 2007; Flint et al., 2005).

1.2.1 Deficits in self-monitoring: Cancellation theory

Various theories have been proposed for the occurrence of hallucinations. Here, I will give a brief overview of one of the most influential accounts of hallucinations, denominated *Cancellation Theory*, that is based on deficits in self-monitoring and a general dysfunction in the ability of identifying self-produced signals as one's own.

Introduction

In 1995, Wolpert and colleagues made a formal proposition of what would become known as the forward model for sensorimotor integration and motor control (Wolpert et al., 1995; further extended in Ghahramani et al., 1997; Miall & Wolpert, 1996; Wolpert & Kawato, 1998). This, theorized that every time the brain produced an action, it would predict the sensory consequences of its own actions (Figure 1.5-A), and learn to attenuate the sensory outcomes of said actions as a way to optimize processing and minimize surprise (Friston & Stephan, 2007). Considering that most actions are goal oriented, this model is ideal for controlling in real-time the outcome of the action, and prepare adequate expectations and reactions (it in fact decreases reaction times; Bubic et al., 2010). It has further been associated with sense of agency (SoA: the subjective experience of being the cause/have control of one's own actions, Haggard & Chambon, 2012), as a result of the comparison between expected and real sensory outcomes. Some have further theorized that by consequence, this can also lead to a distinction between self- and other-produced actions when the prediction of one's action and the actual result do not match (Farrer et al., 2003; Farrer & Frith, 2002).

The Cancellation Theory, proposes that a disruption of this model is at the root of hallucinations. If observed in practice, the forward model acknowledges observations showing that, in healthy individuals, self-produced forces are perceived with less intensity than externally produced ones (Blakemore et al., 1999; Shergill et al., 2003). However, when tested in patients with psychosis, the results did not match those of healthy individuals. Shergill and colleagues (2005) performed a simple but clever experiment with healthy individuals and schizophrenic patients, where they had a lever pressed on their index fingers, and had to replicate (match) this force either by pressing directly on the lever which rested atop of their index fingers, or through a joystick that made the lever move against their fingers (Figure 1.5-B). When pushing directly on the lever to match the force, healthy individuals overdid the amount of force that had been presented before, likely to compensate the sensory attenuation the brain produces on self-generated forces. If the joystick was used (which in the brain's perspective was a mediator with unknown characteristics), then healthy participants could match the force much better. Schizophrenic patients, were also able to match the force well in the joystick condition. What was remarkable, though, was that the patients were better than healthy individuals at matching the force when pressing directly on the lever. This, in the opinion of the authors supported that the attenuation mechanism was not as prominent in patients, as it was in healthy individuals.

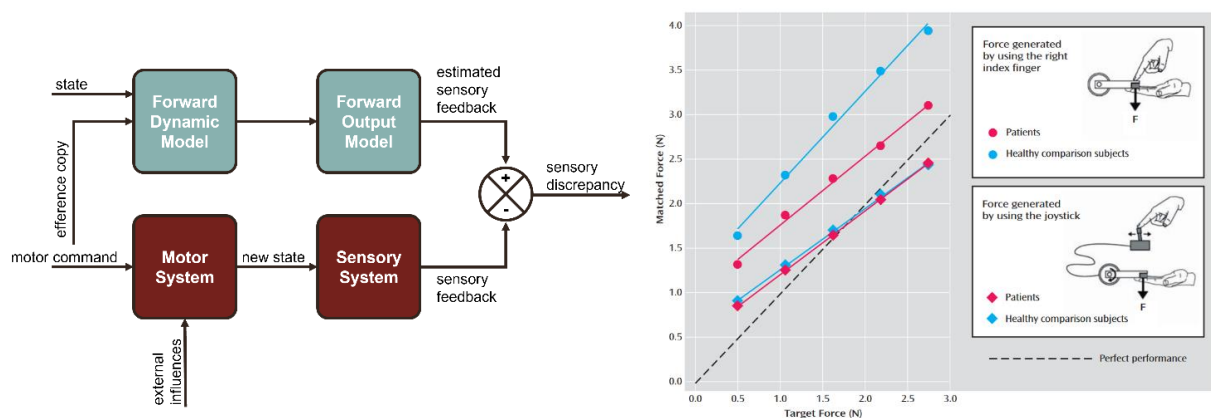


Figure 1.5 Forward models for motor control and consequences in motor action performance

(A) Forward model for motor control. Once a motor command is evoked, an efference copy is produced and fed to a forward dynamic model which uses the efference copy and the current state to recreate the consequences of the motor command. These are then fed to a forward output model which estimates the sensory feedback. The estimated sensory feedback is then compared to the real sensory feedback which comes as a consequence of the actual motor action. A sensory discrepancy is produced. *Based on (Miall and Wolpert, 1996).* **(B)** Results from the matching force experiment. When participants are presented with an external force to their fingers, healthy individuals over estimate the presented force if they have to reproduce it directly on their finger, versus through a joystick. The over-estimation is attenuated in schizophrenic patients, which is believed to be a consequence of impaired self-monitoring. *Reproduced with permission from (Shergill, et al. 2005).*

Introduction

This can also be observed in other domains. Evidence from auditory-verbal hallucinations in patients with schizophrenia supports a failure in attenuating signals, this time related to self-produced speech (Ford et al., 2001, 2002, 2007). Similar findings occur within the visual system but related to somatic passivity (Schnell et al., 2008). Some have argued that the dysfunction of this mechanism is at the (conceptual) “root” of psychosis. Many authors have hypothesized that, in schizophrenia, hallucinations and delusions of alien control are in fact related to failures in recognizing actions as one’s own (Blakemore et al., 2002; Farrer & Franck, 2007; Frith, 1987; Mlakar et al., 1994), and that more broadly, this leads patients to perceive “false saliency” for those actions (for a review: Palaniyappan & Liddle, 2012).

Note that the *Cancellation Theory* is not the only account of psychosis, other exist, such as the Bayesian Theory. In a nutshell, this theory proposes that the brain is as a hierarchical system, which makes sense of noisy surroundings by relying on priors, learned over time, about our contextual environment (de Lange et al., 2018; Kaiser et al., 2019; O’Reilly et al., 2012). Taking this into account, *Bayesians* propose that expected stimuli should carry more weight, and be more salient, than unexpected stimuli, and find evidence on this in neurophysiological studies that show bias in visual representation towards learned representations (Kok et al., 2012, 2013). In relationship to psychosis, *Bayesians* propose that an anomalous encoding of precision in priors or sensory evidence underpins various aspects of psychosis (Adams et al., 2013), with hallucinations being a consequences of highly precise priors (Cassidy et al., 2018; Corlett et al., 2019; O’Callaghan et al., 2017), and delusional beliefs rather a loss of precision (Schmack et al., 2013; Stuke et al., 2018). See Figure 1.6 for an example. At first this theory might be seen as a direct opposite of the *Cancellation Theory*, given it requires higher saliency of expected stimuli. However, the fact that Bayesian theories allow for hierarchy solves this apparent conflict. Namely, the *Opposing Process Theory* proposes that higher weights of priors are seen earlier in the hierarchy (i.e. with regards to brain processing, earlier in time), and cancellation of expected stimuli, later in the hierarchy (i.e. later in time with regards to brain processing). According to this, decisions on agency, saliency, amongst other aspects, are then done based on the distance between the prior and posterior, rather than directly on the weight of the prior, or the difference between expected and real sensory outcome. To see a summary of this reconciliation theory and neurophysiological evidence see (Press et al., 2020).

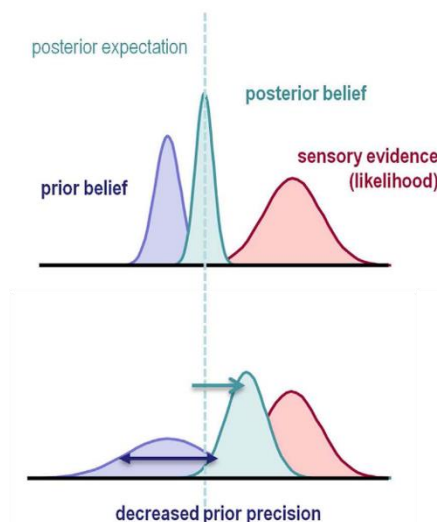


Figure 1.6 Example of Bayesian weighting of prior and sensory evidence

A prior belief and newly available sensory evidence are shown on the panel above, leading to a posterior belief. The panel below then exemplifies how the posterior belief would shift towards sensory evidence, if loss of prior precision occurred. Reproduced with permission from Adams, et al. 2013, under the Creative Commons License 4.0.

1.2.2 Presence Hallucination

The Presence Hallucination (PH), defined as the strange sensation of having someone behind you when no one is actually there, is a central piece to this thesis. It is through this lens, that I studied neural processes of hallucinations, control of hallucinations, and hallucinations in PD, because as we will see, we are able to experimentally induce PH in the lab.

PH is widespread across several different populations. It is considerably present in conditions such as Lewy Body Dementia (Nagahama et al., 2010) and schizophrenia (Llorca et al., 2016), and can also be present, albeit rarely, in epilepsy (Arzy et al., 2006) and stroke (Blanke et al., 2003). Importantly, as will be detailed later on, PH is extremely common in PD (ffytche et al., 2017), can occur before the diagnosis of PD and even before motor symptom onset (Pagonabarraga et al., 2016). Finally, PH can also occur in healthy individuals (Brugger et al., 1999). The case of Reinhold Messner, who experienced PH while doing extreme mountaineering, is an extremely popular example of how PH can happen to healthy individuals (Messner, 2016), and it is often used to illustrate how PH in healthy people tends to occur mostly under extreme circumstances notable examples, e.g. in a shipwreck, see Geiger, 2009).

The phenomenology of PH is highly suggestive of an origin in dysfunctional body and self-processing. Two detailed case studies illustrate this well. The first comes from a 65 years-old nun which presented with seizures and complex neurological disorders (alexia, visual agnosia, hemianopia of the right side), due to a hematoma at the left parieto-temporo-occipital junction (Blanke et al., 2003). Occasionally she experienced PH behind her right side, which importantly was triggered by walking or standing and in a manner that this presence always mimicked her movements. The second case study is that of a woman with drug-resistant epilepsy in whom, while undergoing pre-surgical evaluations of epilepsy, PH could be triggered by intra-cortical electrical stimulation to the temporo-parietal junction (Arzy et al., 2006). In her case, the presence closely mirrored all of her movements, appeared in seemingly impossible physically locations (e.g. while laying in bed, as if the presence was merged with the bed), and seemed to interfere with experimental tasks. These two cases studies illustrate phenomenological aspect of mimicking of body positions and performed actions, which is highly suggestive of a relationship between PH and bodily processing. Although I have chosen to give these illustrative examples, literature has similar descriptions in large cohorts of patients with different diseases (Brugger et al., 1996) and in PD (Fenelon et al., 2011). Recently we have also published a paper on two case studies of PD patients whose PH was exclusively triggered by locomotor activities such as walking (Potheegadoo et al., 2022; see appendix).

In sum, these clinical and phenomenological observations point to a disturbance of bodily processing being at the origin of PH. In the next section, I will elaborate on this potential root of PH, in the light of the topics introduced in the previous chapter, and in the light of bodily-self consciousness – a minimal theory of self-consciousness, where the body takes on a major role (Park & Blanke, 2019).

1.2.2.1 A disruption in bodily-self consciousness

Self-consciousness is the experience of identifying ourselves with an “I”, a *self*. In our lab, we have argued that in its simplest form, this experience is rooted in multisensory integration of bodily signals, and is referred to as bodily self-consciousness (BSC; Park & Blanke, 2019). In its main features, BSC generates a minimal selfhood through: self-identification, this body belongs to *me* and *I* identify myself with it; self-location, knowing where *I* am; and first-person perspective, from where *I* perceive the world (Blanke, 2012).

Introduction

Important for the understanding of PH is the core component of BSC, self-identification. Different avenues of research have shown that self-identification can be experimentally manipulated. For example, through touches to a participant's real hand (hidden from sight) and visually-congruent touches to a rubber hand placed in a congruent anatomical position, the rubber-hand illusion has participants identifying the rubber hand as their real hand (Botvinick & Cohen, 1998; Ehrsson et al., 2004, 2007). At the whole body level, this has been done through the full-body illusion (Lenggenhager et al., 2007) or the body swap illusion (Petkova & Ehrsson, 2008), which again make use of coherence between visuo-tactile stimuli, with an added change in perspective. In neuropsychiatric and healthy populations the dysfunction of core components of BSC can lead to Autoscopic Phenomena (AP), which are illusory reduplications of one's own body (Blanke et al., 2004; Blanke & Mohr, 2005; Brugger et al., 1997, p. 199). These have been associated with dysfunction at the level of the temporo-parietal junction (TPJ), a hub for multisensory body processing (Blanke & Arzy, 2005; Blondiaux et al., 2021).

With regards to the PH, it is possible that a dysfunction at the core of self-identification is contributing to its rise. Considering that PH does not include any of the typical sensory components such as vision and audition (Brugger et al., 1999), but is rooted in sensorimotor processing as seen in the previous chapter, it is possible that PH is a reduplication of one's own body, which is perceived as a foreign agent (Peter Brugger et al., 1996). In this sense, PH as a reduplication arising from dysfunctional sensorimotor processing has been mostly considered in the light of the *Cancellation Theories* (Blakemore et al., 1998, 2002), with bodily signals not being appropriately attenuated and leading to the perception of an external agent, which is in fact one's doppelgänger (Bernasconi et al., 2021; Blanke et al., 2014; Case et al., 2020).

1.2.2.2 Experimental induction of Presence Hallucinations

Inspired by the clinical observations made in patients with PH, the theories behind it, and developments in neuroscience-robotics, Blanke and colleagues (2014) developed a robotic system capable of inducing PH in controlled laboratory settings. In this setup, a participant stands between the two parts of the robotic system while blindfolded and isolated from surrounding noise. The participant then manipulates the front part of the robotic system by moving an extended pulley in a poking like fashion, while the back part mimics their movements and provides tactile feedback to the participant's back, either synchronously or asynchronously (Figure 1.7). It is in this asynchronous stimulation that the PH is typically induced (Bernasconi et al., 2022; Blanke et al., 2014).

In a previous subsection (1.2.1), I introduced *Cancellation Theories* for psychosis, and it is in fact likely that this plays a significant role in this induction. As participants move their arm to produce poking-like movements and provide tactile feedback to their own backs, the mechanism previously described for prediction of sensory outcomes is at play (it is at play for any action we do). Both the spatial and temporal conflicts are likely to blur these predictions, however of the two, the temporal aspect is likely to have a much more prominent effect. Spatial conflicts are present in the real world, for example if one uses a tool to scratch a part of their body, and can be modelled by the brain if the tool is known, or upon a few usages of the tool. Temporal delays are more salient to the brain. This, because the brain is hypothesized to have a *temporal binding window* in which multisensory stimuli, can be integrated and interpreted together (Colonius & Diederich, 2004; van Wassenhove et al., 2007). That window is hypothesized to extend to approximately 250 milliseconds, and has a normal distribution for likelihood of integration. Hence stimuli that is presented at its edges can lead to erroneous perceptions (for example see the induction of visual illusion through sound: Shams et al., 2002). Here, we hypothesize that receiving the touch on the back in a delayed manner, is interfering

Introduction

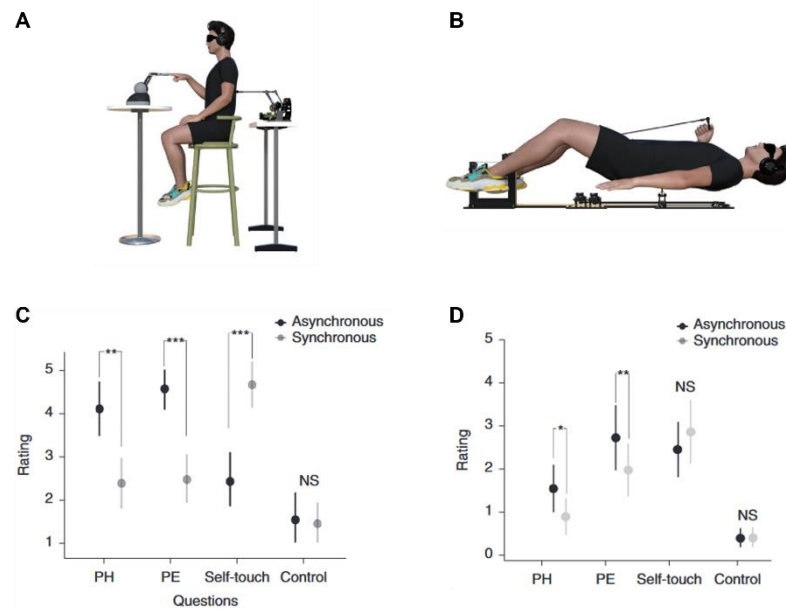


Figure 1.7 Robotic setup to induce PH and induction results

(A) Robot equipment to induce PH. A front-part of the robotic system allows participants to generate poking-like movement that are then reproduced on their backs by the back-robot, either in the asynchronous or synchronous condition (500 msec of delay, and no delay, respectively). **(B)** MRI-compatible version of the robot. **(C)** Results of the questionnaire measuring PH, accompanying sensations (PE, self-touch), and control questions. As seen, PH is induced with stronger intensity in the asynchronous conditions. **(D)** Results for the MRI-compatible robot. Adapted with permission from (Bernasconi, et al. 2022).

with the typical multisensory integration process, and causing the back touch, to not be integrated as the real sensory outcome of the produced movement, but rather as a salient event, potentially attributed to an external source.

Several studies have by now induced PH in healthy individuals using this method, and many have studied the effect of PH-induction – and by extension of hallucinations – on several cognitive functions. The original study using riPH, reported an interesting effect of numerosity judgment of people in the same room. When entering the experimental room participants would see four people inside a room and were told that at any given point these people could approach them and they would be asked to judge the number of people close to them while still being blindfolded. Participants significantly overjudged the amount of people close to them in the asynchronous condition, that typically induces PH, even though the four people would promptly leave the room after the experiment started (Blanke et al., 2014). In another experiment, the induction of PH was shown to blur the barrier between self- and other-generated thoughts across a series of experiments that relied on source monitoring, but not in an experiment that did not (Serino et al., 2021; see appendix). Another study showed that sensorimotor conflicts and the induction of passivity experiences (PE; the sensation that someone other than yourself is producing your own actions) through this robotic system, increased the perception of quiet sounds, but not in people who experienced strong PE (Orepic et al., 2021). Also investigating auditory perception but in early psychosis patients, a different study found that these patients were sensitive to the induction of PH, and the condition with maximum sensorimotor conflicts modulated their auditory perception in a self/other distinction task (Salomon et al., 2020).

Introduction

Crucially, we have argued that such a system presents an incredible opportunity to study hallucinations directly in the lab (Bernasconi et al., 2022; see appendix).

1.2.2.3 A tool to study hallucinations in clinical populations

Due to their erratic nature, hallucinations are quite complicated to study in the lab, as they tend to occur during one's everyday life. Hence, there have been different ways to try and circumvent this problem. Studies can resort to the use of psycho-active drugs to study hallucinations, however, these typically induce hallucinations across many modalities, with uncontrollable intensity, and can furthermore be accompanied by considerable changes in consciousness (Baggott et al., 2010). Others have developed methods to induce visual hallucinations, such as Ganzfeld- (Wackermann et al., 2008) and flicker-induced hallucinations (Allefeld et al., 2011), however, these are phenomenologically incomparable to what the patients feel in real life and lack appropriate control conditions. Furthermore, both these methods and drug-induced hallucinations, fail to induce hallucinations in repetition. Another possibility is to study false-positives in a task where sensory cues are presented around the threshold of perception (e.g. Vercammen et al., 2008). While this is an interesting approach, it is arguable whether a false-positive necessarily constitutes a hallucination. Finally another approach and probably the most typical is to stratify patients in different groups to then study the neural correlates of hallucinations by focusing on the group that has them (e.g. Chang et al., 2017). While this is a valid avenue of research that has allowed us to significant progress in knowledge, it could be argued, that it is rather identifying a trait (that potentially makes subjects more prone to hallucinate), but not the state itself, which is conducive of the hallucination.

In this sense, the robotic system developed in Blanke's group is extremely promising, as it allows for the study of hallucinations directly in the lab (Bernasconi et al., 2022; Blanke et al., 2014). The most notable example in our case is that of PD. When the robotic system is used in patients with PD, it is much more likely to trigger PH in patients that experience it in daily life, as compared to patients that do not (Bernasconi et al., 2021). This allows us and other researchers to effectively study the PH, as it is induced.

1.2.3 Parkinson's Disease

PD is a neurodegenerative disease marked by degeneration of nigrostriatal dopaminergic neurons (Postuma & Berg, 2016), and is characterized by a series of motor-symptoms including, but not limited to, tremors, bradykinesia (difficulty in initiating movements), dyskinesia (uncontrolled movements), stooped posture, and a limitation in facial expressions typically known as masked face (Kalia & Lang, 2015; Postuma & Berg, 2016). Albeit being more known for its motor symptoms, PD is accompanied by prominent and early non-motor symptoms which include sleep disorders, depression, cognitive decline, and frequently hallucinations (Aarsland et al., 2021; Hely et al., 2008; Ravina et al., 2007).

As mentioned previously, hallucinations and in particular visual hallucinations, are quite prevalent in PD, affecting up to 70% of patients (ffytche et al., 2017; Hely et al., 2008). Descriptions of these hallucinations became more common after the introduction of dopaminergic medication (Diederich et al., 2009), however reports of such hallucinations are known before the introduction of this medication (Fenelon et al., 2006). It has become clear that in the past, they were severely underreported in PD patients due to the prejudices associated with them (Wood et al., 2015). This was particularly exacerbated by the fact that PD patients retain insight in their hallucinations for a considerable duration of disease (Diederich et al., 2003) and there

Introduction

was no appropriate screening by clinicians (Chan & Rossor, 2002). Hence, authors generally concur that hallucinations are an intrinsic part of the underlying pathology of PD, even if they might be exacerbated by dopaminergic medication (Diederich et al., 2009; ffytche et al., 2017).

1.2.3.1 Minor Hallucinations

In addition to visual hallucinations, PD patients are commonly afflicted by a sub-group of so-called *minor hallucinations* (Fenelon, 2000). These include presence hallucinations, the convincing sensation of having someone close by when no one is actually there, passage hallucinations, the false sensation of having something/someone pass by you, and visual illusions which include pareidolias, misperceptions of movements, amongst other types of illusions. Interestingly, these minor hallucinations can occur before the diagnosis of PD in up to 40% of patients, and might precede the occurrence of motor symptoms in up to 15% of patients (Pagonabarraga et al., 2016)[†].

The development of minor hallucinations in PD is now of considerable interest in research, as they typically progress into more complex hallucinations – including visual hallucinations – and might also progress into loss of insight and delusional ideations (ffytche et al., 2017; Goetz et al., 2006[‡]; Lenka et al., 2019). In that sense, understanding the neurophysiology and brain dynamics behind minor hallucinations, might hold great value for PD patients, as it could unlock an understanding of how to act early in the disease course, and prevent further progression.

1.2.3.2 Progression to structured hallucinations and cognitive decline

The progressive nature of hallucinations in PD, is well established. However, the consequences of this progression might not only be limited to hallucinations. The occurrence of visual hallucinations and persistent psychosis in PD, has been extensively linked to progressive cognitive decline in longitudinal studies (Factor et al., 2003; Hely et al., 2008). This relationship is however not straightforward. While some have hypothesized that hallucinations might potentiate cognitive decline (Ibarretxe-Bilbao et al., 2010; Morgante et al., 2012; Ramirez-Ruiz et al., 2007), others have presented evidence for the key role of cognitive decline in loss of insight over the hallucinations (Llebaria et al., 2010), and the particular role of frontal-lobe dysfunction in PD associated psychosis (Lenka et al., 2017).

Importantly one of the aims of this thesis was to further investigate the links between minor hallucinations, hallucinations and cognitive decline (Figure 1.8). Recently, minor hallucinations have been hypothesized to mark increased neurodegeneration associated with subjective (but not objective) cognitive decline, potentially indicating a worse form of the disease (Bejr-kasem et al., 2021). Hence, given that complex hallucinations are linked to cognitive decline and that minor hallucinations might progress to complex hallucinations, in this thesis I have used stratified cohorts of patients to study if minor hallucinations might set off an early/more severe forms cognitive decline.

[†] The reader should consider recall bias as an important limitation to this study's last claim

[‡] The term benign hallucinations used in Goetz, et al. (2006) did not refer only to minor hallucinations, but to hallucinations with retained insight.

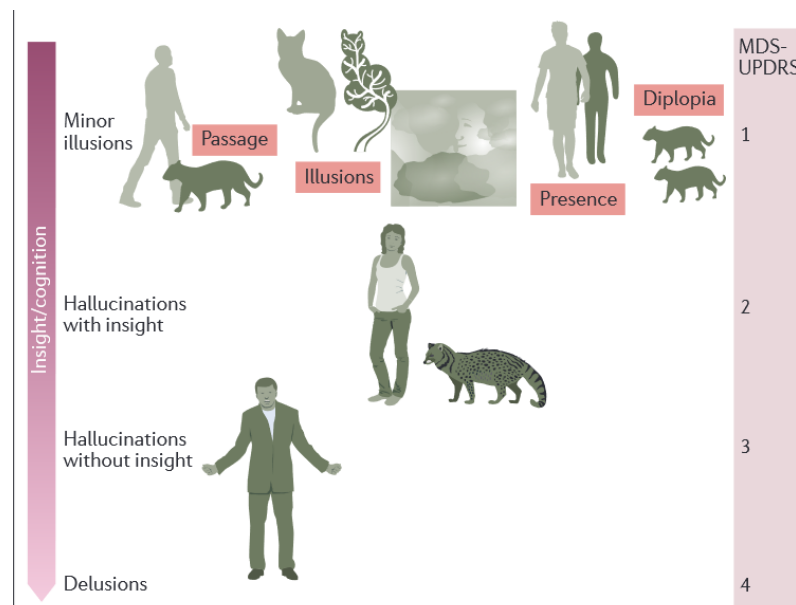


Figure 1.8 Progression of hallucinations and insight/cognition

A proposed hierarchy and progression of neuropsychiatric symptoms following the progression of the disease (indicated in the MDS-UPDRS column). As significant evidence ties minor hallucinations to hallucinations, and hallucinations to persistent psychosis and cognitive decline, it is possible that minor hallucinations are already an indicator of a worse form of the disease. Adapted with permission from (ffytche, et al. 2017).

1.3 Neuroimaging of Hallucinations

In this chapter, I will now summarize some of the main findings with regards to neuroimaging theories and findings on hallucinations, that were key to the development of this thesis.

1.3.1 Evidence for dysconnectivity in hallucinations

When recording brain activity from patients with psychosis and hallucinations, a consistent finding across different studies is the altered connectivity across brain regions (for a review: Stephan, et al., 2009). This dysconnectivity tends to be found between frontal and temporo-parietal regions (Baker et al., 2014; Bluhm et al., 2007; Skudlarski et al., 2010), and within the frontal cortex (Cole et al., 2011), with the latter being associated with both cognitive dysfunction and positive symptoms[§] in psychosis (Camchong et al., 2011; Fornito et al., 2011). In similar fashion, widespread dysconnectivity is reported across multiple functional brain networks, such as the DMN, CEN, and SN (Karbasforoushan & Woodward, 2012; Woodward et al., 2011; Yu et al., 2015). The SN exhibits reduced intra-connectivity in schizophrenic patients (White et al., 2010), which was linked to an aberrant salience processing in such patients (Pankow et al., 2015) and further linked

[§] Positive symptoms within the psychosis spectrum typically refer to hallucinations, delusions, and some changes in behavior, whereas negative symptoms refer to depression and withdrawal.

Introduction

to the severity of hallucinations (Smieskova et al., 2015). Other studies have also identified dysfunctional intra-connectivity in both the DMN (Camchong et al., 2011; Whitfield-Gabrieli & Ford, 2012) and CEN (Fornito et al., 2011), in schizophrenic patients.

It has been argued that the specific reduction in fronto-parietal connectivity seen in these patients might account for a dysfunctional role of prediction signals (as discussed in the previous chapter) and lead to the misattribution of an external agent and generation of hallucinations (Friston et al., 2016). However interesting and informative this hypothesis is, I would argue that dysconnectivity over the fronto-parietal or temporo-parietal regions is not enough to explain hallucinations. Conceptually it would be strange if a dysfunction, believed to be constant in psychosis, was to be responsible for symptoms that fluctuate over time. Moreover, these findings tend to not always have the same direction of dysfunction (sometimes hyper-connected, other times hypo-connected, across the same regions; Stephan, et al., 2009). Hence, these findings are rather likely to be a trait aspect of the psychosis spectrum, rather than a state which can explain hallucinations. With this in mind, we will explore in a much more fine-grained manner, the temporal processes of brain functioning associated with hallucinations and psychopathology.

1.3.2 Dynamic brain processes

In 2011, Vinod Menon published an comprehensive review on a process he believed could underpin hallucinations in general (Menon, 2011). It highlighted a “triple network model” focused on hierarchical interactions of the DMN, CEN, and SN (Figure 1.9) – in which switching between the DMN and CEN is controlled by the SN depending on context, attentional demand, and mostly salience (Goulden et al., 2014) – that Menon believed to be dysfunctional in psychopathology. In particular, Menon proposed that aberrant saliency mapping could be at the root of several symptoms of psychosis, including hallucinations.

Saliency processing, an essential aspect of the triple network model, is known to be impaired in the psychosis spectrum (White et al., 2010). In particular, the insula, a crucial node of the SN when it comes to saliency detection (Menon & Uddin, 2010), presents with major dysfunctions in sensitivity to subjective saliency in psychosis patients (Uddin, 2015). Using dFC methods able to decompose the underlying time-structure of brain activity, Palaniyappan and colleagues (2013), identified that the areas such as the parahippocampus and the dlPFC do not exert as much influence on the insula in schizophrenic patients, compared to healthy individuals. This might be particularly relevant, for erroneous misattribution of actions, considering the role of the dlPFC in action selection and monitoring (Chambon et al., 2013; Nahab et al., 2011). In turn, the right anterior insula has been shown to have reduced activity (Manoliu et al., 2014), and less able to orient the balance between CEN and DMN (measured as decreased lag correlation between SN to DMN and CEN, but increased between the latter two), when comparing schizophrenic patients (during psychosis and outside psychosis) with healthy people (Manoliu et al., 2014; Moran et al., 2013). Another recent study in a large cohort of schizophrenic patients showed that reduced and less occurring network interactions between the triple-network system could identify 80% of the patients, and was significantly correlated to their hallucinatory symptoms (Supekar et al., 2019). Finally, a study which acquired fMRI during hallucination experiences of schizophrenic patients, used dynamic causal modelling (DCM) to model the relationships within the triple network system and between this system and the hippocampus (Lefebvre et al., 2016). Researchers argued that the initiation of hallucinations are marked by a destabilization of the triple-network system, which makes it more vulnerable to inappropriate hippocampal stimuli during the actual hallucinations (for more information on hippocampal role in hallucinations see: Cachia et al., 2020; Jardri et al., 2011). In other words, the SN deemed hippocampal stimuli relevant and forced a decrease in activity of

Introduction

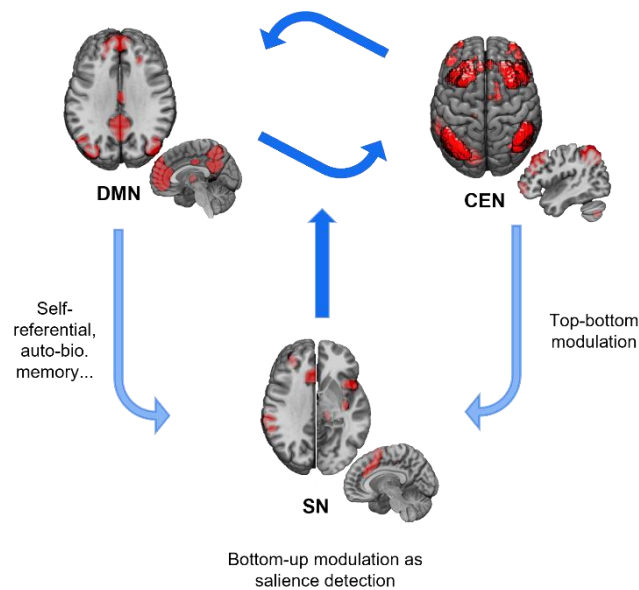


Figure 1.9 Triple Network Model

Organization of the Salience Network (SN), Central Executive Network (CEN), and Default-Mode Network (DMN) according to the triple network model. The SN modulates the balance between the CEN and DMN, based on contextual saliency. In turn, the DMN also exhibits influences on the SN based on self-referential processes and auto-biographical memories, and the CEN also exerts influence on the SN as top-down modulation.

DMN, which the authors argue is abnormal for rest. Hallucinations then come to a stop when the system is capable of engaging the CEN enough, to direct cognitive attention.

The findings highlighted here, have underlined the major relevance of three aspects with regards to hallucinations. One, salience processing. Two, balance across large-scale networks, in particular within the triple-network model. And three, temporal processes are crucial for hallucinations. These findings were in their majority found across modalities, speaking for a common general mechanism underlying cross-modality hallucinatory experiences. Later, we will focus on the mechanisms underlying a specific hallucination, however a more interested reader might wonder how these *general* hallucination processes relate to hallucinations in specific modalities. Evidence suggests that as the DMN disengages during the period where hallucinations are occurring, the associative cortices, but not the primary ones, engage, with which associative cortex depending on the modality of the hallucination (Jardri et al., 2013).

1.3.3 Dynamic brain processes in Parkinson's Disease

As mentioned previously, hallucinations are a significant non-motor symptom of PD. Here, I will focus on the studies which have targeted dynamic connectivity across large-scale networks.

The aberrant salience processing theory mentioned in the previous chapter has also been theorized for PD. However, in this case, a focus was put on poor recruitment of the dorsal attentional network (DAN) as being at the root of hallucinations, which in turn leads to an increased dependency on the DMN during ambiguous perceptions (Shine et al., 2011). It is important to note that the formal distinction is made at the level of the executive network itself not being able to activate or be recruited, whereas before the failure was in the attribution of saliency (and hence a failure of the SN), which lead to an imbalance between the DMN and CEN.

Introduction

Using a task designed to induce higher number of visual misperceptions in PD patients with visual hallucinations versus those without, researchers observed a link between increased misperceptions and decreased activity of the DAN during the task (Shine et al., 2014). Interestingly, they also found that the recruitment of the insula was crucial to task performance, and that participants with lower performance exhibited lower functional connectivity between the DAN and the SN during resting-state fMRI. This provided some evidence, that the dysfunction was not just at the level of the CEN, but to an extent also a failure of recruiting the DAN by part of the SN (note however that this link was only done through resting-state data). Later studies analyzed temporal dynamics of these networks in these patient cohorts. Visual misperceptions were associated with higher activity of the SN and DMN, but crucially, more misperceptions were linked to a decrease in network connectivity between a visual network and the DAN and a higher network connectivity between the same visual network and the CEN (Figure 1.10; Shine et al., 2015). With this, a significant role of the DMN was established for hallucinations in PD, together with a failure in activating the CEN. Finally, in a recent study researchers found that those with visual hallucinations presented with a detriment in the ability to remain in specific brain states, and transition with increased ease across different brain states (Zarkali et al., 2022). An intrinsic facilitation of transitions across these states could explain the prominence of the DMN and its maladaptive increase of interactions with sensory networks.

Not as much is known for minor hallucinations. One single study on the matter, albeit not on dynamic connectivity, has found significant increases in connectivity within the DMN and between the DMN and poles of an executive network (Bejr-kasem et al., 2019). In this thesis, I also studied neural correlates of hallucinations across a patient cohort stratified in minor and complex hallucinations.

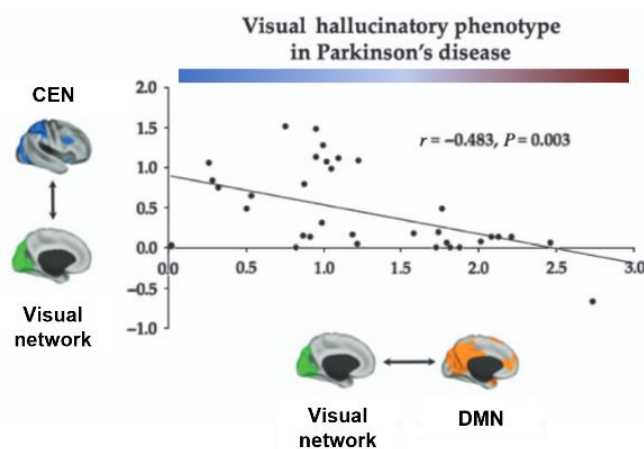


Figure 1.10 Network interactions between the DMN, CEN and a visual network in hallucinations in PD

As patients increase in their hallucinatory phenotype (blue to red), the visual network is seen decoupling from the CEN, and rather progressively couple more with the DMN. Adapted with permission from (Shine, et al. 2015), under the Creative Commons License 4.0.

1.3.4 Neuroimaging evidence related to the Presence Hallucination

The PH is a specific hallucination, and hence only a few studies have investigated its particular mechanisms. Dysfunction at the level of the TPJ has been shown for PH in three separate studies (two already presented in detail in chapter 1.2.2 Presence Hallucination). One study presented evidence of a hematoma at the level of the TPJ, in a patient experiencing PH (Blanke et al., 2003). Another showed that focal electrical stimulation in the posterior superior temporal sulcus (pSTS) could induce a strong PH in an epileptic patient (Arzy et al.,

Introduction

2006). Finally, a third study which performed lesion overlap analysis in 12 patients with neurological PH, observed considerable overlap at the level of the TPJ, but also in the frontoparietal cortex and the insula (Blanke et al., 2014).

Studies using fMRI targeting PH, have also been possible, as the robotic system used to induce PH described in subsection “1.2.2.2 Experimental Induction of Presence Hallucinations” has also been extended to MRI settings (Hara et al., 2014), and was used to study PH induction in healthy individuals inside the MRI (Bernasconi et al., 2021; Blondiaux, 2020). In this study participants manipulated the robot in the synchronous and asynchronous conditions for 30 seconds, interleaved by periods of 15 seconds of rest. A cluster of four regions, the posterior superior temporal sulcus (pSTS), inferior frontal gyrus (IFG), anterior insula (aINS), and medial superior frontal gyrus (mSFG), were found to be more active in the asynchronous versus synchronous condition, hence being specific to the temporal conflict. Another, much larger cluster but not overlapping, was tied to the spatial conflict (participants perform movements in the front but receive feedback on their backs). In the same study, the authors used lesion network mapping analysis (Boes et al., 2015) to identify functional networks affected by neurological lesions of patient suffering from PH, but without PD. The overlap of this lesion network with the two mentioned above identified three regions, the pSTS, IFG, and ventral premotor cortex (vPMC), which were related to the PH-induction task in healthy individuals and PH in neurological patients. When studied in PD patients, this group of regions had significant disconnectivity, and remarkably, the disconnection between the pSTS and IFG could predict with an accuracy up to 97% whether patients with PD had PH or not (Bernasconi et al., 2021). Significant disconnection between these two regions was also found in psychosis patients suffering from PE as compared to patients without PE (Stripeikyte et al., 2021).

One of the goals of this thesis was to identify the neural correlates of PH induction, given that this was, previously, not possible. Namely, the neural processes that underlie hallucinations which are arguably *hidden* in the temporal dimension, and can only be found with methods that explore temporal properties. This will be undertaken in Part I – Study 1. Another objective of this thesis was to identify the correlates of minor hallucinations specifically in PD patients, through the use of PH-induction. Significant effort was dedicated to having PD patients manipulate the MR-compatible robot in the scanner and conduct our task there. Unfortunately, this was not possible, due to the motor comorbidities of our patients who were most sensitive to the induction. While solutions are procured to make this task feasible for our patients inside the MRI, the work addressed in Part II – Study 3, studying the neural correlates of minor hallucinations and PH in PD was done mostly through the use of our PH-induction task outside the scanner as a proxy for resting-state differences across groups.

1.4 Real-time fMRI neurofeedback

One of the main objectives of this thesis was to achieve volitional control of the neural underpinnings of hallucinations, in an attempt to modulate hallucinations themselves. To do so, given that we have the capability to induce PH inside the MR-scanner (Bernasconi et al., 2021, 2022), we turned to neurofeedback (NF) using real-time fMRI.

The concept behind neurofeedback is relatively simple: the same way one can learn a skill, a sport, or covertly learn associations, self-regulation of brain activity is possible as long as some sort of *informative feedback* is provided (Sitaram et al., 2017). By pairing NF with fMRI (Figure 1.11), we can achieve this goal of providing

Introduction

online feedback of brain activity, and hence allow one to enact control over their own brain activity (Caria et al., 2012; Sulzer, Haller, et al., 2013; Weiskopf et al., 2007; Weiskopf, 2012).

Since its proof-of-concept (deCharms et al., 2004; Yoo & Jolesz, 2002), real-time fMRI NF has been used in a range of applications. The first was to reduce chronic pain (deCharms et al., 2005). Participants learned to control the rostral anterior cingulate cortex, and managed to decrease pain intensity, in a manner that was sustained post-training and out-performed autonomic feedback. Other applications have been seen for example for attention (deBettencourt et al., 2015), cognitive performance (Megumi et al., 2015; Yamashita et al., 2017), memory (Sherwood et al., 2016), fear-exposure (Koizumi et al., 2017), attention-deficit/hyperactivity disorder (Arns et al., 2014), and motor symptoms in PD (Subramanian et al., 2011), amongst many others (for an overview of different applications see: Thibault et al., 2018). For the purpose of this thesis, I was mostly interested in using real-time fMRI NF to regulate brain neural correlates of hallucinatory states in healthy individuals, to pave the way for future clinical translation.

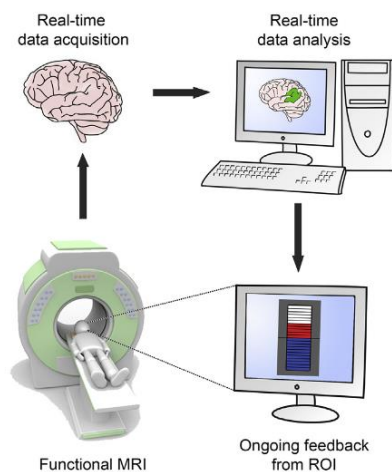


Figure 1.11 Real-time fMRI neurofeedback

Example of a real-time fMRI neurofeedback setup in which data is processed in real-time from the ongoing fMRI acquisition, and used to present a visual stimuli to the participant, indicating the activation of a particular brain region of interest (ROI). The participant is then expected to develop a mental strategy to regulate the ROI. Adapted with permission from Thibault, et al 2018.

1.4.1 Neuropsychiatric applications of real-time fMRI neurofeedback

Real-time fMRI NF has been used for a variety of neuropsychiatric conditions (for a review see: Yamada et al., 2017). Its potential to target functional biomarkers, characteristic to many of these conditions, has made it particularly interesting, especially with the advent of real-time fMRI neurofeedback based on static and dynamic functional connectivity metrics (Watanabe et al., 2017). In this thesis, I focus on its application to hallucinations. Orlov and colleagues (2018) were one of the first to show that a cohort of patients suffering from schizophrenia and with auditory-verbal hallucinations (AVH) could achieve volitional control of the superior temporal gyrus (STG). While this in itself did not change the severity of AVH (and no control condition was present), a link was found between decreased severity and increase in connectivity between the STG and IFG, a posteriori. Later on, by focusing directly on this specific functional connectivity (posterior STG and IFG), researchers were able to have schizophrenic patients strengthen this connection, which resulted in a subjective decrease in the perceived burden of the disease, but not specifically AVH (Zweerings et al., 2019). Another study attempted to decrease AVH by focusing on the anticorrelation between the DMN and CEN (Bauer et al., 2020), an important feature of brain activity considering the triple network model theory of hallucinations (Menon, 2011; introduced in Chapter "1.3.2 Dynamic brain processes"). By using an explicit NF strategy where patients were told they should do a specific meditation strategy, an increase in anti-correlation between the DMN and CEN was observed, together with a decrease in severity of AVH. However, the lack of an appropriate well-matched control condition, meant it was not possible to know if meditation

or NF lead to these changes, especially when considering that meditation itself can lead to reduction in the burden of AVH (Sheng et al., 2019).

1.4.2 Methodological considerations in real-time fMRI neurofeedback

With the widespread use of this technique, various differences in methodology have arise, mainly across feedback presentation and on how to control for confounding factors. Feedback can be presented either in a continuous or intermittent manner, with the former providing more opportunities for learning and the latter a more precise type of feedback (Stoeckel et al., 2014). Research has also shown that the participant does not need to be aware of what the feedback is related to (behaviorally) for it to have a regulatory effect, making a case for implicit learning in real-time fMRI NF (Shibata et al., 2011). Controlling for confounds is typical across the majority of real-time fMRI experiments. Depending on the goals of the experiment, this might be done using a comparison against gold-standard treatment/intervention, or by having participants do mental-rehearsal controls for potential behavioral effects. Additionally, controlling for placebo effects can be done through the use of a different group, exposed to sham feedback or feedback from a brain region of no interest. However, the “holy-grail” of real-time fMRI NF control is the bidirectional regulation paradigm, as it controls within-participant for all the above mentioned confounds and can establish the neuro-specificity of the biomarker with respect to the behavioral goal (Sorger et al., 2019). Another way to do so is by using an adaptive type of paradigm, in which the feedback is embedded in the task itself, hence also controlling for these different confounds within participant (deBettencourt et al., 2015).

Another aspect in which there are significant differences across real-time fMRI NF experiment is the target of NF training. Initially, most of the paradigms targeted activation of a single brain region (Caria et al., 2007; Sulzer, Sitaram, et al., 2013). However, with technological development that allowed for faster processing, different methodologies that would otherwise require long processing times, were brought to the scanner. For example, using functional connectivity (e.g: Spetter et al., 2017; Zilverstand et al., 2014), multivariate pattern analysis-like approaches, such as decoded NF (e.g: Amano et al., 2016; Shibata et al., 2011), DCM-based neurofeedback (Koush et al., 2013), and recently, full-network NF, has also been tested (Bauer et al., 2020; Pamplona et al., 2020).

Naturally, the target of NF training should be adequate to the targeted biomarker, and so should the design of the experiment. In the present thesis, I have opted for full network NF, given that it is the biomarker of induced PH and that it holds great promise in providing a translation from a successful and specific regulation to behavioral changes (Bauer et al., 2020; Watanabe et al., 2017; Yamada et al., 2017). Importantly, my task design also included a within-participant baseline, which ultimately allowed for bidirectional regulation.

1.5 Thesis at a glance

The main goal of this thesis was to identify the neural correlates of PH induction, investigate whether such neural correlates are prone to regulation, and if so, whether their successful regulation could modulate PH. The second goal of this thesis was to identify distinct neural correlates of minor hallucinations across a stratified cohort of PD patients with different hallucinations, and to relate such neural correlates to the sensitivity to PH induction using our robotic-system, as well as to clinical signs of cognitive impairment. Finally, the third goal, achieved in between the first two, was to elucidate the optimal setup for riPH experiments, motivated by the translation of this research from healthy to patient populations.

Hence, **Part I**, incorporates two studies, and is dedicated to the first goal. With **Study 1** I have made use of the concepts of dFC (Preti et al., 2017) to explore the neural correlates of PH, from a dynamical perspective of brain activity. Following this, I then designed a real-time fMRI NF paradigm (**Study 2**) to have participants achieve volitional control of the identified neural correlates, while PH is being induced. This study pioneered in the methodology of control, as it is the first to our knowledge to do so concomitantly with hallucination induction.

In **Part II**, I leveraged our experimental setup to induce PH in PD patients, to assess sensitivity to induction of PH in these patients across a cohort with distinct hallucinations (**Study 3**). This, alongside with neuropsychological assessments, was then used together with resting-state fMRI, to investigate neural correlates associated to minor hallucinations and cognitive decline. Effectively, this attempts to build the foundation for a future real-time fMRI NF study targeting the neural correlates identified here.

Part III represents the third goal of this thesis which was performed in between the first two major goals. Here, I used Bayesian methods for meta-analysis, to quantify the exact effects of PH-induction, and advise on the optimal setup (**Study 4**).

Finally, this is then followed up by a general discussion putting these findings in perspective, and an appendix which includes other work for which I have contributed during my doctoral studies (this work fits within my research area but not necessarily with the main goals of my thesis).

1.6 Overview and personal contributions

Personal contributions are labelled below for each part using the CRediT taxonomy (Allen et al., 2019).

Part I: Identifying and modulating neural correlates of PH

Study 1: Robotically-induced hallucination triggers subtle changes in brain network transitions

Dhanis, H., Blondiaux, E., Bolton, T., Faivre, N., Rognini, G., Van De Ville, D.[§], Blanke, O.[§] (2022). *NeuroImage* 248, 118862

Personal contribution: Conceptualization (idea and formulation), methodology, software, investigation (data acquisition), formal analysis, resources, data curation, writing – original draft, writing – review and editing, visualization.

Study II: Sustained bidirectional self-regulation of hallucination networks through real-time fMRI neurofeedback

Dhanis, H., Gninenko, N., Morgenroth, E., Potheegadoo, J., Rognini, G., Faivre, N., Blanke, O.[§], Van De Ville, D.[§]. *In preparation*

Personal contribution: Conceptualization (idea and formulation), methodology (development of specific methodology on top of existing one), software (significant adaptation of OpenNFT, development of experiment specific software), validation, investigation (data acquisition), formal analysis, resources, data curation, writing – original draft, writing – review and editing, visualization.

Part II:

Study 3: Robot-induced hallucinations and PLS-CAPs reveal neural correlates of preserved cognitive and psychosis-like traits in PD

Dhanis, H., Potheegadoo, J., Bernasconi, F., Stampacchia, S., Maradan M., Stucker, Horvath, J., Fleury, V., Wicki, B., Bally, J., Ghika, J-A, Burkhard, P., Krack, P., Van De Ville, D., Blanke, O.

Personal contribution: Conceptualization (formulation), methodology, software, investigation (data acquisition), formal analysis, data curation, writing – original draft, writing – review and editing, visualization.

Part III: Quantifying and optimizing PH induction

Study III: Mega-analysis of presence hallucination induction experiments using robotically mediated sensorimotor conflicts

Dhanis, H., Potheegadoo, J., Faivre, N., Bernasconi, F., Blanke, O. *In preparation*

Personal contribution: Conceptualization (idea and formulation), methodology, formal analysis, data curation, writing – original draft, writing – review and editing, visualization.

APPENDIX

Supplementary study 1:

Bolton, T., Tuleasca, C., Wotruba, D., Rey, G., **Dhanis, H.**, Gauthier, B., Delavari, F., Morgenroth, E., Gaviria, J., Blondiaux, E., Smigielski, L., Van De Ville, D. (2020). TbCAPs: a toolbox for co-activation pattern analysis. *NeuroImage*, 211, 116621

Personal contribution: validation

Supplementary study 2:

Serino, A., Pozeg, P., Bernasconi, F., Solcà, M., Hara, M., Progin, P., Stripeikyte, G., **Dhanis, H.**, Salomon, R., Bleuler, H., Rognini, G., Blanke, O. (2021). Thought consciousness and source monitoring depend on robotically controlled sensorimotor conflicts and illusory states. *iScience*, 24(1), 101955.

Personal contribution: Investigation (data collection), software (for carrying out specific experiments in the study)

Supplementary study 3:

Potheegadoo, J., **Dhanis, H.**, Horvath, J., Burkhard, P. R., & Blanke, O. (2022). Presence hallucinations during locomotion in patients with Parkinson's disease. *Movement disorders clinical practice*, 9(1), 127.

Personal contribution: Conceptualization (organization), formal analysis (review and critique), writing – review & editing.

Supplementary study 4:

Bernasconi, F.*, Blondiaux, E.*, Rognini, G., **Dhanis, H.**, Jenni, L., Potheegadoo, J., Hara, M., & Blanke, O. (2022). Neuroscience robotics for controlled induction and real-time assessment of hallucinations. *Nature Protocols*, 1-24.

Personal contribution: Software (development of the GUI software for behavioral experiments)

* authors contributed equally

Part I: Identifying and modulating neural correlates of PH

2.1 Study 1: Robotically-induced hallucination triggers subtle changes in network transitions

Herberto Dhanis ^{1,2,3,4}, Eva Blondiaux ^{1,2}, Thomas Bolton ^{3,4}, Nathan Faivre ^{1,2,5}, Giulio Rognini ^{1,2}, Dimitri Van De Ville ^{1,3,4, §,*} and Olaf Blanke ^{1,2,6, §,*}

Affiliations

¹ Center for Neuroprosthetics, Ecole Polytechnique Fédérale de Lausanne, EPFL, Geneva, Switzerland

² Brain Mind Institute, Faculty of Life Sciences, Ecole Polytechnique Fédérale de Lausanne, Lausanne, Switzerland

³ Institute of Bioengineering, Ecole Polytechnique Fédérale de Lausanne, Lausanne Switzerland

⁴ Department of Radiology and Medical Informatics, University of Geneva, Geneva Switzerland

⁵ University. Grenoble Alpes, University Savoie Mont Blanc, CNRS, LPNC, 38000 Grenoble, France

⁶ Department of Clinical Neurosciences, University Hospital of Geneva, Geneva, Switzerland

[§] *Both authors contributed equally*

^{*} *Corresponding author*

Corresponding authors

Prof. Olaf Blanke

Bertarelli Chair in Cognitive Neuroprosthetics
Center for Neuroprosthetics & Brain Mind Institute
School of Life Sciences
Ecole Polytechnique Fédérale de Lausanne (EPFL)
Campus Biotech, H4.3
Ch. des Mines 9
CH-1202 Geneva
E-mail: olaf.blanke@epfl.ch
Tel: +41 (0)21 693 69 21

Prof. Dimitri Van De Ville

Center for Neuroprosthetics & Institute of
Bioengineering
School of Engineering
Ecole Polytechnique Fédérale de Lausanne (EPFL)
Campus Biotech, H4.3
Ch. des Mines 9
CH-1202 Genève
E-mail: dimitri.vandeville@epfl.ch

2.1.1 Abstract

The perception that someone is nearby, although nobody can be seen or heard, is called presence hallucination (PH). Being a frequent hallucination in patients with Parkinson's disease, it has been argued to be indicative of a more severe and rapidly advancing form of the disease, associated with psychosis and cognitive decline. PH may also occur in healthy individuals and has recently been experimentally induced, in a controlled manner during fMRI, using MR-compatible robotics and sensorimotor stimulation. Previous neuroimaging correlates of such robot-induced PH, based on conventional time-averaged fMRI analysis, identified altered activity in the posterior superior temporal sulcus and inferior frontal gyrus in healthy individuals. However, no link with the strength of the robot-induced PH was observed, and such activations were also associated with other sensations induced by robotic stimulation. Here we leverage recent advances in time-resolved analysis, which have been applied to different psychiatric conditions, to decompose fMRI data during PH-induction into a set of co-activation patterns that are tracked over time, as to characterize their occupancies, durations, and transitions. Our results reveal that, when PH is induced, the identified brain patterns significantly and selectively increase their transition probabilities towards a specific brain pattern, centred on the posterior superior temporal sulcus, angular gyrus, dorso-lateral prefrontal cortex, and middle prefrontal cortex. This change is not observed in any other control conditions, nor is it observed in association with other sensations induced by robotic stimulation. The present findings describe the neural mechanisms of PH in healthy individuals and identify a specific disruption of the dynamics of network interactions, extending previously reported network dysfunctions in psychotic patients with hallucinations to an induced robot-controlled specific hallucination in healthy individuals.

Keywords

Presence Hallucination; Dynamic Functional Connectivity; Co-activation Pattern Analysis; Network Interactions; Psychosis; Robotics

2.1.2 Introduction

The sense of presence or presence hallucination (PH) is the sensation of feeling another person close by when in fact no one is actually there (James, 1902). It has been described as an incomplete hallucination, which although vividly perceived, cannot be attributed to any of the usual “sensible ways”, such as visual and auditory perception (James, 1902; Jaspers, 1913). PH has been reported in a variety of medical disorders ranging from stroke (Blanke et al., 2003) to epilepsy (Blanke et al., 2014), brain stimulation during invasive presurgical evaluations (Arzy et al., 2006a), and schizophrenia (Jaspers, 1913). PH is also one of the most frequent hallucinations in Parkinson’s Disease (Diederich et al., 2009; Fénelon et al., 2011) and has also been reported by healthy individuals in extreme situations (i.e., Messner, 2016).

Clinical evidence suggests that altered processing of bodily and sensorimotor signals is an important mechanism in PH, given the ‘sharing’ of posture, position, and movement between the patient and the ‘presence’, as well as the association of PH with sensorimotor deficits (Brugger et al., 1996, 1997; Blanke et al., 2008). Although the paroxysmal and short-lasting characteristics of PH made it difficult to study this hallucination, Arzy and colleagues (2006a) demonstrated that the PH can be induced repeatedly and in a controlled fashion through electrical brain stimulation of the temporo-parietal junction (TPJ), a major integration hub for multisensory and sensorimotor bodily signals (Matsushashi et al., 2004; Arzy et al., 2006b; Blanke, 2012). Lesion mapping in a group of neurological patients with PH suffering from focal brain damage confirmed the importance of the TPJ, and also revealed contributions of insular and frontoparietal cortices (Blanke et al., 2014).

Based on these clinical data, Blanke and colleagues (2014) showed that a robotic setup, that allows to expose participants to different conditions of sensorimotor stimulation, is capable of inducing a sensation comparable to PH (albeit of lower intensity). By providing tactile stimulation to the participants’ own backs, the intensity of such robotically induced PH can be controlled by either changing the force feedback or the delay between the movements and the touch feedback. Thus, the PH occurs more often if there is a delay between the movements the participant is performing with the front part of the robotic system and the touch feedback given by the back part of the robotic system on the participant’s back (asynchronous condition; Figure 1). The PH is absent or of lower intensity when movement and feedback occur at the same time (synchronous condition). Using MR-compatible robotics and fMRI in healthy participants, an extended network was identified to be associated with PH (Bernasconi and Blondiaux et al., in press). Moreover, these authors used lesion network mapping analysis from neurological patients reporting symptomatic PH to further corroborate the PH network, leading to three areas that overlapped with the regions revealed by fMRI during robot-induced PH: the posterior part of the middle temporal gyrus and superior temporal sulcus (pSTS), ventral premotor cortex (vPMC), and the inferior frontal gyrus (IFG) (Bernasconi and Blondiaux et al., in press). From these three regions, only the pSTS and IFG differed in their activity between the asynchronous versus synchronous condition and were further considered for the present study.

Despite the implication of pSTS and IFG regions in the difference between asynchronous versus synchronous condition, their activities did not correlate with the intensity of the robot-induced PH. In addition, concomitant to the induction of the PH in the asynchronous condition, robotic stimulation also induced certain passivity experiences (PE; i.e., the sensation that someone else is touching your body; Mlakar et al., 1994), with most participants that experienced PH reporting it in unison with PE,

while others that did not experience PH still reporting PE. Considering that clinical observations have highlighted the paroxysmal nature and short duration of PH (Blanke et al., 2008; Fénelon et al., 2011), we hypothesize that the temporal dynamics of the PH's neural underpinnings might share these aspects, and hence could be revealed with methods detecting dynamic changes in brain activity.

In the present study, we set to identify the neural mechanisms of the PH in more detail. We focus particularly on studying the temporal dynamics and more subtle changes in brain activity that might underlie PH. To do so, we investigated fMRI BOLD signal during the robotic sensorimotor task, used to induce the PH (Bernasconi and Blondiaux et al., in press), and applied recently established dynamic functional connectivity (dFC) methods that can capture whole-brain network fluctuations in short time ranges (Hutchison et al., 2013; Preti et al., 2017) and that have shown promising results in the study and differentiation of psychiatric conditions (Damaraju et al., 2014; Rashid et al., 2014; Bolton et al., 2020a). Specifically, we apply Co-Activation Patterns (CAPs; Liu and Duyn, 2013) analysis to investigate the dynamically occurring and spatially distributed activity patterns that reflect functional networks associated with the induction of PH and of PE. CAP analysis is based on the assumption that when the BOLD signal is high in relevant seed regions (Tagliazucchi et al., 2011, 2012), different CAPs are expressed at different moments in time (Liu and Duyn, 2013). As seeds, we chose the two key regions (pSTS, IFG) that were associated with both the PH-inducing condition in healthy people and the network in neurological patients with PH (Bernasconi and Blondiaux et al. in press). The CAPs related to these seeds were characterized by their occupancy, average duration, and transition probabilities (Chen et al. 2015; Bolton et al., 2020b), and compared across the two experimental conditions and rest, as well as between different intensities of PH and PE.

We hypothesized that the neural mechanisms of PH will be grounded in a temporary dominance of certain CAPs, which would either have increased occupancy, average duration, or a shift in transition probabilities favouring one or more CAPs, only during the induction of PH, but not of PE. Furthermore, we predicted that these mechanisms of PH will differ from those of PE, in terms of brain anatomy, recruiting different brain regions related to the asynchronous delay condition (Bernasconi and Blondiaux et al., in press).

2.1.3 Materials and methods

The present study performed dFC analysis on the data from Bernasconi and Blondiaux, and colleagues (in press). The following sections will summarise the participants included, and the experimental design used in that study, as well as the analysis and methodologies employed in the present study.

2.1.3.1 Participants

25 healthy individuals (10 females) with a mean age of 24.68 (\pm 3.70, range 18-32) years old took part in the PH-induction experiment (study 2.1 in Bernasconi and Blondiaux et al., in press). Every participant was right-handed as assessed by the Edinburgh Handedness Inventory (Oldfield, 1971). All participants gave their informed consent prior to the start of the experiment, following the Declaration of Helsinki, and the study was approved by the local ethics committee of the Canton of Genève, Switzerland.

2.1.3.2 Experimental Paradigm

Throughout the experiment, participants were blindfolded with an eye mask and wore both ear protection and headphones, in an effort to maximally isolate them from the surroundings. Laying on top of a special platform-bed, that concealed the back part of the robotic system used to induce PH, participants could manipulate a lever attached to the front part of the robotic system (figure 2.1A). The robot itself, composed of a front and back part, allows its users to provide tactile feedback on their own backs. This is achieved by moving the front part of the system with a lever, which controls the back part of the robot that provides tactile feedback on the participant's back. A conflict in the spatial domain is hence always present, with the movements performed in the front space being perceived immediately on the back space (synchronous condition). A second conflict can be introduced in the temporal domain, by delaying the feedback received on the back (asynchronous condition). In the asynchronous condition where these two conflicts are combined, the sensation of having someone behind you (PH), and the sensation that someone else producing is your actions (PE), can be elicited in healthy individuals (Blanke et al., 2014).

The task itself consisted of 16 blocks of 30 seconds of robot manipulation, interleaved by blocks of 15 seconds of rest (figure 2.1B). Two conditions were assessed: the synchronous condition in which the movements performed by the participants with the front part of the system were synchronously reproduced onto their backs, and the asynchronous condition where a delay of 500 ms was introduced between the performed movement and the tactile feedback. The conditions were presented randomly to the participants, with no same condition being delivered more than twice in a row. In total, each participant performed two runs, with a total duration of 25 minutes.

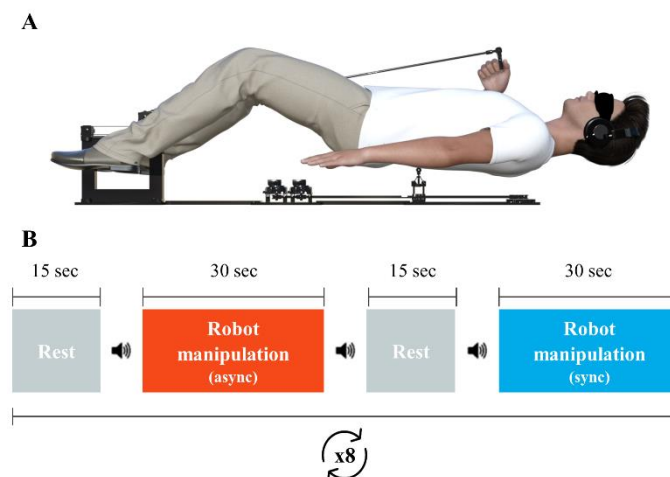


Figure 2.1 Robotic system and experimental paradigm

(A) MR-compatible robotic system used to induce the PH in healthy individuals. The robotic system is composed of a front part, which has a manipulator stick slide on a rail, and allows the participants to move in the x (bottom and up) and z (up and down) directions. These movements are then transmitted to a back robot that is confined below an MR-compatible platform. In the present experiment the back robot mimics the movements of the front part of the robotic system either in real-time, or with a delay of 500 milliseconds. **(B)** During the experiment participants perform two runs, each with sixteen 30 second blocks of robotic manipulation, interleaved with 15 seconds of rest. The blocks of robotic manipulation are performed in the synchronous condition, where movements with the front robot are reproduced in real-time by the back robot, or in the asynchronous condition, where movements with the back robot are reproduced by the back robot with a delay of 500 milliseconds. Blocks of the same type do not appear more than twice in a row. Auditory cues passed on to the participant through headphones, mark the beginning and ending of each block. (Adapted from Bernasconi and Blondiaux et al., in press).

Part I: Identifying and modulating neural correlates of PH

At the end of the scanning session, participants performed 30 seconds of robotic manipulation for each condition (i.e., synchronous and asynchronous) in a counterbalanced fashion. After each condition, participants answered a questionnaire assessing their subjective experience during the robotic stimulation (7-point Likert scale, see Supplementary Table S1). The questionnaire included questions such as, “I felt as if I was touching my body”, to assess self-touch impressions, “I felt as if someone else was touching my body”, to assess PE, and “I felt as if someone was behind me” to assess PH. For each participant we computed the strength of the induced sensations, respectively for each questionnaire item, as the difference between the score in the asynchronous and synchronous conditions. A positive score reflects an induced sensation that is stronger in the asynchronous condition, whereas a negative score indicates a stronger sensation in the synchronous condition.

2.1.3.3 MRI data acquisition

Functional image acquisitions were performed at the MRI facility of the Campus Biotech (Geneva, Switzerland), with a Siemens MAGNETOM Prisma 3T scanner, and using a 64-channel head-and-neck coil. For the sensorimotor task, echo-planar sequences were used (EPI, TR = 2.5s, TE = 30 ms, with a flip angle of 90°, GRAPPA = 2), with a resolution of 2.5x2.5 mm, and a slice thickness of 2.5 mm (no gap, 43 slices). Anatomical images were acquired with T1-weighted MPAGE sequences (192 slices, FOV = 240 mm, TR = 2.3s, TE = 2.32ms).

2.1.3.4 Dynamic Functional Connectivity Analysis through Co-Activation Patterns

CAPs analysis is based on point process analysis (Tagliazucchi et al., 2012) and temporal clustering (Liu and Duyn, 2013). In particular, given the activity time course of one or more seed regions, different dynamically occurring network configurations that co-activate with these seeds are extracted. For the present implementation, we built upon the TbCAPs toolbox (Bolton et al., 2020b). The data was pre-processed using custom MATLAB (MATLAB 2019b) scripts and SPM12 functions (Wellcome Department of Cognitive Neurology, Institute of Neurology, UCL, London, UK). Volumes were realigned to the first scan, and spatially smoothed with a gaussian filter (FWHM = 6mm), after being normalized to MNI space. Then, linear trends were removed from each voxel's time course, and such time courses were further temporally z-scored.

The first step of CAPs analysis requires the selection of one or more seed regions in order to identify timepoints when these seeds exhibit high BOLD signal. As seeds we considered the right pSTS and right IFG, which are the only two regions that were previously reported to be more active in the asynchronous condition (where PH and PE are induced) than in the synchronous condition, and are part of a functionally impaired network in neurological patients experiencing symptomatic PH (Bernasconi and Blondiaux et al. in press). In the second step of the analysis, timepoints of any of the two seed regions where activity exceeded a z-score of 1, were marked and considered for further analysis. To deal with head motion, timepoints with a framewise displacement above 0.5 mm were scrubbed (Power et al., 2012). In the third step, the volumes (frames) of the selected timepoints were fed into a k-means algorithm to obtain temporal clusters based on spatial patterns. The best k was selected beforehand through consensus clustering which provides stability measures for data points

being clustered together across different numbers of selected centroids (Monti et al., 2003). In the fourth and final step, all frames assigned to the same label were averaged to obtain a representative CAP. The frames of timepoints when none of the seeds were active are averaged in a non-active state, CAP0. Finally, each timepoint is tagged, taking into account haemodynamic lag of 2 TRs, to one of the three experimental conditions: asynchronous sensorimotor manipulation, synchronous sensorimotor manipulation, or rest.

2.1.3.4.1 Occupancy and average duration of the CAPs

For each CAP (CAP1 to CAP9) and for the non-active state (CAP0), two metrics were computed during the different experimental conditions: occupancy and average duration (Chen, et al., 2015). Let $CAP_i[k]$ be the binary time course that indicates if CAP i , is active at any timepoint k element of the set $D_{S,C}$ which contains all active, non-active, and scrubbed, timepoints of a condition C , for a single participant S . Occupancy refers to the percentage of scans a CAP i occupies in a given condition:

$$Occ(i; D_{S,C}) = \frac{\sum_{k \in D_{S,C}} CAP_i[k]}{|D_{S,C}|}$$

We define $CAPDur_i[r]$ as the duration (number of consecutive timepoints active) of CAP i for each associated occurrence r . Average duration is the mean CAP duration for each participant:

$$AvgDur(i; D_S) = \overline{CAPDur_i(r)}$$

With the main goal of identifying specific CAP behaviour in the asynchronous versus synchronous condition, as well as its association with PH and PE, CAPs' occupancies were tested across conditions by means of a non-parametric ANOVA (Friedman's test). When significant effects were observed, this was followed by post-hoc non-parametric tests of the medians (i.e., Wilcoxon rank sum test) for each of the CAPs to reveal significant changes in occupancy with the experimental different conditions. These results were corrected for the false discovery rate (FDR) with the Benjamini-Hochberg procedure (Benjamini and Hochberg, 1995) to correct for multiple comparisons (i.e., number of CAPs).

The average duration of a CAP was only computed for the actual occurrences of a CAP. If a CAP never occurred for a participant in a specific condition, that measure of average duration was not considered zero, but rather the participant was excluded for the assessment of that specific measure. As consequence the number of measures of average duration, per CAP and per condition, was not always equal to the number of participants. These different sample sizes meant that average duration could not be tested with a Friedman's test. We hence, tested average duration with Linear Mixed Models (LMM), using the `lmer` function provided with the package `lme4` (Bates et al., 2015) available for R (version 3.6.1). The average duration was modelled as a function of CAP, condition, interaction between CAP and condition, and a random-intercept accounting for inter-participant variability.

$$AvgDur \sim (1|Participant) + CAP + Condition + CAP:Condition$$

Part I: Identifying and modulating neural correlates of PH

The parameters of the model were tested for significance by sequential comparison of the simplest model with only the random-participant intercept against a model adding the CAP parameter, followed by comparing the latter model with one adding the condition parameter, and finally comparing this one with the full model that includes the interaction parameter. Comparisons were performed with the Wald χ^2 test (Liu, 2016). If an interaction was detected, the effect of condition on average duration was then investigated for each CAP.

If a CAP was found to have higher occupancy or average duration in the asynchronous condition, Spearman's correlations were used to investigate potential relationships between that CAP's occupancy or average duration and the strength of the subjective experiences of PH and PE.

2.1.3.4.2 Transition probabilities between the CAPs

To characterize temporal relationships between CAPs, we computed transition probabilities (TPs) that describe the probability of a CAP to transition to itself, to another CAP, or to the non-active state (CAP_0). Per condition and per participant, the TPs of an initial CAP i to a next CAP j are computed by normalising the number of times a starting CAP i transitions to a next CAP j , by the number of times the initial CAP i occurs:

$$TP(i, j; D_S) = \frac{\sum_{k \in D_S} CAP_i[k] CAP_j[k+1]}{\sum_{k \in D_S} CAP_i[k]}$$

The TPs can then be organized per participant and condition into a 10x10 matrix, which will be considered as the TP matrix characterizing a first-order Markov chain modelling the sequence of CAPs.

The goal was now to investigate if the Markov model changes, under different experimental conditions, and between the subgroups of participants who are sensitive (or not) to PH and/or PE induction. To that aim, we modelled the TPs with LMM. Our approach to this problem can be described as a three-step hierarchical analysis of the factors that can influence TPs.

We first modelled the TPs based on the initial and next CAPs involved in each transition. This implied modelling the data with a fixed-effect parameter for the initial CAP, a fixed-effect parameter for the next CAP, and an interaction parameter for the initial and next CAPs. Between-participant variability is accounted for in the model with a random intercept:

$$TP \sim (1 | Participant) + CAP_{init} + CAP_{next} + CAP_{init}:CAP_{next}$$

I
II
III

We assessed how each parameter improved the explained variance of the data, by consecutive comparisons of increasingly complex models (0, I, II, III) with a Wald

Part I: Identifying and modulating neural correlates of PH

χ^2 test. Model 0 only comprised a random-effect parameter for between participant variability. Model I included in addition, a fixed-effect parameter for the initial CAP (i.e., row in the TP matrix). Model II added a fixed-effect parameter for the next CAP (i.e., column in the TP matrix). And Model III added the interaction between the initial and the next CAP, which expresses that TPs for each pair of CAPs can be different.

In case Model III revealed a significant effect of the interaction parameter, we continued with the second step of the analysis. We divided the subsequent analysis into two parts. First, fixing a specific initial CAP (i.e., row of the TP matrix), and, second, fixing a specific next CAP (i.e., column of the TP matrix). For clarity, we will refer to them as analysis on the forward and backward properties of the Markov models, respectively. The corresponding LMMs are represented as follows, with F or B representing the forward or backward properties, and i the fixed initial or next CAP:

Forward Properties $TP(CAP_{init} = CAP_i) \sim (1|Participant) + CAP_{next} + Cond + PH + Cond:PH$

F_i 0

F_i I

F_i II (PH)

F_i III (PH)

Backward Properties $TP(CAP_{next} = CAP_i) \sim (1|Participant) + CAP_{init} + Cond + PH + Cond:PH$

B_i 0

B_i I

B_i II (PH)

B_i III (PH)

These models included the effects of, condition, sensitivity to PH induction (positive PH-strength score), and interaction between these two. The sensitivity to the induction of the PH can be interpreted as a group variable as it is a binary value. To assess the effect of sensitivity to PE induction (positive PE-strength score) and interaction with condition, another set of models was used (for F_i and B_i , Models II and III). This was due to convergence not being achieved with the current amount of data available if both PH, PE, and respective interactions with conditions, are analysed simultaneously. To compensate for this, we lowered the significance threshold of the subsequent Wald χ^2 tests to 0.025, instead of 0.05. As before, the effect of each model's parameters was assessed by sequentially comparing simpler models and more complex models, e.g. from Model B_i 0 to Model B_i III (PH), using the Wald χ^2 test.

Finally, if a significant effect of the interaction between condition and sensitivity to PH-induction or to PE-induction, was observed for a specific B_i III Model, then that model continued into the third and last step of the analysis. Here, to properly investigate the effect of sensitivity to PH or PE induction on the TPs over different conditions, the models also fixed the conditions (these models will have C added in

the notation, as in $F_i^c(\text{PE})$ with C taking the values a , s or r , identifying the asynchronous, synchronous or rest condition, respectively). An example for investigating the effect of sensitivity to PH-induction, on the TPs of a specific $CAP_{next} = CAP_i$ in the synchronous condition, is the following model:

$$TP(CAP_{next} = CAP_i, Cond = Sync) \sim (1|Participant) + CAP_{next} + PH$$

$B_i^s 0 \text{ (PH)}$
$B_i^s 1 \text{ (PH)}$

2.1.4 Results

2.1.4.1 CAP analysis reveals multiple distinct spatial brain patterns

With CAP analysis, we explored how different CAPs occurred and interacted under different experimental conditions, and, secondly, how this was associated with the level of the subjective experience of PH and PE. The identified CAPs consisted of brain regions co-activating with one or both seeds, during periods in which at least one of the seeds was active (z-score above 1). The IFG seed was considered active on average 14.7% (SD: ± 2.9) of the total timepoints of the entire experiment (i.e., all sequential periods of rest, robot manipulation in the asynchronous condition, robot manipulation in the synchronous condition; in their original order for each participant, respectively). A similar percentage of 15.0% (SD: ± 2.1) was found to be active for the pSTS seed. On average, 23.6% (SD: ± 3.4) of all timepoints were selected, given that seed activity overlapped at given times. The remaining timepoints were assigned to the non-active state, with only 0.4% (SD: ± 1.0 %) being scrubbed. Following timepoint selection, consensus clustering was run to determine the best number of centroids for the clustering procedure. The stability measure assessed with this method recommended segregating the data into 9 centroids (Supplementary Figure S1). Hence, CAPs analysis was applied to the selected timepoints revealing nine different co-activation patterns (Supplementary Figure S2 and Table S2, for description of cluster peaks), for which occupancy, average duration, and TPs were explored (see below).

2.1.4.2 Sensorimotor conditions alter CAPs occupancy and average duration

Here, we will first focus on the analysis of a CAP's occupancy (percentage of a CAP occurrence in a condition) and a CAP's average duration (number of seconds a CAP lasts on average, once it occurs) across the two sensorimotor conditions and rest.

Occupancy. Friedman's test revealed a significant difference in occupancy of CAPs between the conditions of asynchrony, synchrony, and rest (p-value = 0.020). Follow-up multiple comparison corrected post-hoc tests revealed that several CAPs changed their occupancy depending on the condition (Figure 2.2A). CAP 6 was the only CAP to show a higher occupancy in the asynchronous condition, when compared to both the synchronous condition (p-value = 0.049) and to rest (p-value = 0.002). This CAP's increase in occupancy between the asynchronous and synchronous condition did not show any significant correlation with the strength of the subjective experiences of PH ($p = 0.10$, p-value = 0.630) or with PE ($p = 0.23$, p-value = 0.280). CAP 7 and CAP 9 had a significantly higher occupancy for both asynchronous

Part I: Identifying and modulating neural correlates of PH

and synchronous conditions, when compared to rest (both p -values < 0.001). However, the occupancy of these two CAPs did not differ between the asynchronous and synchronous conditions. A summary of the occupancy results for all remaining CAPs can be seen in the Supplementary Table S3.

Average duration. Linear mixed models fixed-effect statistics revealed a significant effect of CAP on the average duration (p -value < 0.001), showing that different CAPs have different average durations. A significant effect of condition was also identified (p -value < 0.001), indicating that the experimental conditions significantly changed the average durations of the CAPs. Crucially, a significant interaction between CAP and condition was observed (p -value < 0.001), showing that the experimental conditions affected the average duration of each CAP differently. To further investigate this interaction, we ran post-hoc tests for the effect of condition on each CAP (Figure 2.2B). CAP 6 showed a significant difference in average duration across the conditions, with the asynchronous condition having a higher average duration than the synchronous condition and rest (estimate in asynchronous condition: 3.56, SE = ± 0.17 , t -value = 20.66; estimate effect of synchronous condition: -0.60, SE = ± 0.24 , t -value = -2.55; estimate effect of rest: -0.57, SE = ± 0.24 , t -value = -2.33; p -value = 0.021). This increase of average duration did not show any correlation with the strength of the induced PH ($\rho = 0.06$, p -value = 0.78), nor that of PE ($\rho = 0.05$, p -value = 0.78). CAP 7 also showed a significant difference in average duration across the conditions, with the asynchronous condition having a higher average duration than the synchronous condition and rest, and the synchronous condition also lasting longer than rest (estimate in asynchronous condition: 3.41, SE = ± 0.16 ; t -value = 21.81 estimate effect of synchronous condition: -0.30, SE = ± 0.22 , t -value = -1.40; estimate effect of rest: -0.64, SE = ± 0.22 , t -value = -2.95; p -value = 0.015). This increase of

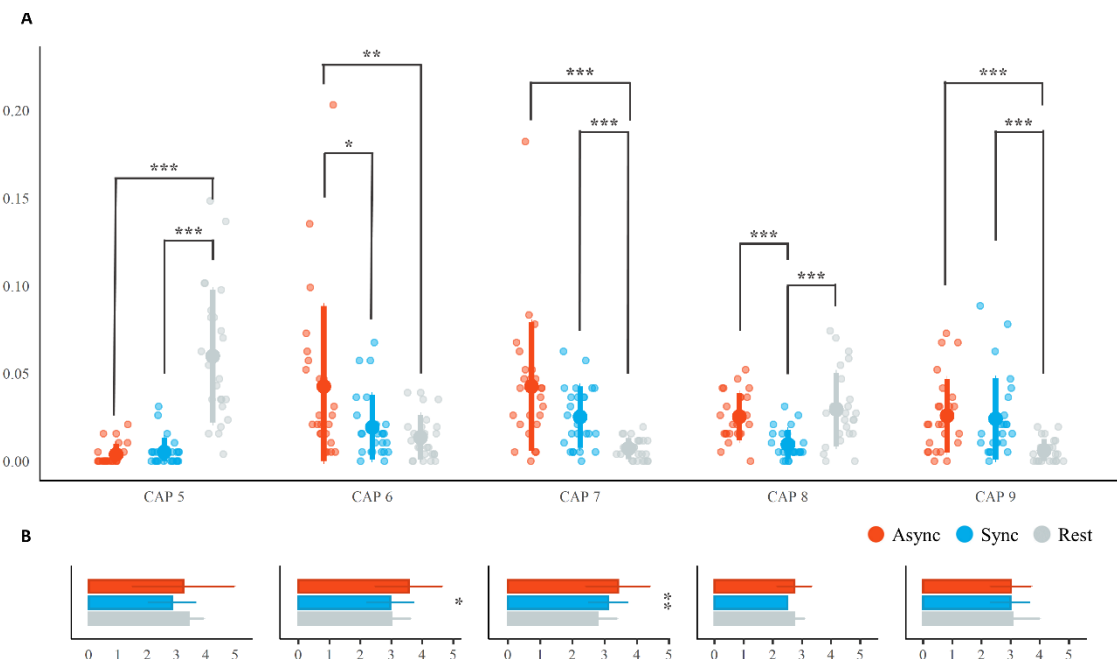


Figure 2.2 Occupancy and average duration of the CAPs

(A) The occupancy of CAPs 5 to 9 in the different conditions. While CAPs 7 and 9 were associated with the sensorimotor conditions, CAP 6 showed a significantly higher occupancy for the asynchronous condition as compared to synchrony (and rest), denoting its specificity to the temporal conflict present in the asynchronous condition. **(B)** The average duration of CAPs 5 to 9. Only CAPs 6 and 7 show an effect of condition on the average duration, with CAP 6 lasting more in the asynchronous condition, and CAP 7 lasting more, the more sensorimotor conflicts are introduced. (* p -value < 0.05 ; ** p -value < 0.01 ; *** p -value < 0.001).

Part I: Identifying and modulating neural correlates of PH

average duration for CAP 7 did not show any correlations with the strength of induced PH ($p = 0.04$, $p\text{-value} = 0.84$), nor that of induced PE ($p = 0.12$, $p\text{-value} = 0.55$). The values for all other CAPs were not characterized by significant results ($p\text{-value} > 0.05$) and are summarised in Supplementary Table S4.

Brain activation for Co-Activation Pattern 6 The network identified as CAP 6 (figure 2.3A) is composed by ten brain regions, with its main components in the right pSTS, bilateral inferior parietal lobule (IPL, centred on the angular gyrus, AG), the right dorso-lateral prefrontal cortex (dlPFC), which included the IFG region used as seed, and the middle pre-frontal cortex (mPFC; including part of the supplemental motor area, SMA). Smaller cortical regions were found in the left dlPFC and left premotor cortex. Subcortical activations were detected in the right caudate and the cerebellum (left Crus I and II). Deactivations were observed in the cuneus and the occipital gyrus.

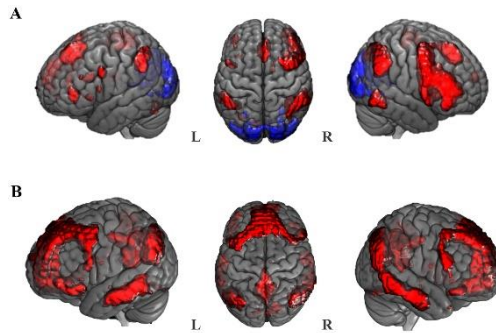


Figure 2.3 Anatomy of CAP 6 (PH) and CAP 9 (PE)

(A) Brain regions of CAP 6 are shown involving the right posterior superior temporal sulcus, the bilateral inferior parietal lobule with focus on the angular gyri, the right dorso-lateral prefrontal cortex, the middle prefrontal cortex (including part of the supplemental motor area), the left dorso-lateral prefrontal cortex, the left precentral gyrus, the body of the caudate on the right, and in the cerebellum, the left Crus I and II. Deactivations are observed over the cuneus and occipital gyrus. **(B)** Brain regions of CAP 9 are shown in posterior cingulate cortex, middle prefrontal cortex and bilateral posterior parietal cortex. This CAP extends from the DMN with clusters over the superior and middle frontal gyrus, bilateral clusters on the middle temporal gyrus, and bilateral clusters on the Crus I and II of the cerebellum. Two small clusters are also observed in the thalamus.

2.1.4.3 Experience of PH changes CAP's transition probabilities

Next, we investigated to what extent CAPs' transitions are affected by the experimental conditions (asynchronous, synchronous, rest) and by sensitivity to the induction of PH and PE (more information on the model's results in supplementary table S5).

In the first step of this analysis, we assessed whether TPs depended on the initial and next CAP, by comparing the first three models. Linear mixed models fixed-effect statistics revealed significant effects for all the parameters associated with transitions between CAPs. Specifically, Model I revealed a significant effect of the initial CAP ($p\text{-val} < 0.001$), showing that CAPs' TPs will change depending on the initial CAP of the transition. Model II revealed a significant effect of the next CAP ($p\text{-val} < 0.001$), showing that CAPs' TPs also depended on which CAP they are transitioning to. Finally, Model III revealed a significant interaction effect between the initial CAP and next CAP of a transition ($p\text{-val} < 0.001$), showing that TPs are dependent on the specific combination of initial and next CAP. Given this interaction effect, we proceeded with the analysis of the remaining parameters of condition, PH sensitivity, and

Part I: Identifying and modulating neural correlates of PH

PE sensitivity, and did so separately for each value of initial (forward properties) and next (backward properties) CAP.

Forward properties. Here we focused on analysing the remaining parameters mentioned above, for each fixed initial CAP level ($CAP_{init} = CAP_i$). By doing so, Models F_i I revealed a significant effect of condition (figure 2.4A, 2.4B, 2.4C, Supplementary Figure S3) for transitions beginning in CAP_{init} 5 (p-value < 0.001) and beginning in CAP_{init} 9 (p-value = 0.003), as seen in figure 2.4D. The remaining models, F_i II (PH) and F_i III (PH), showed that PH sensitivity did not

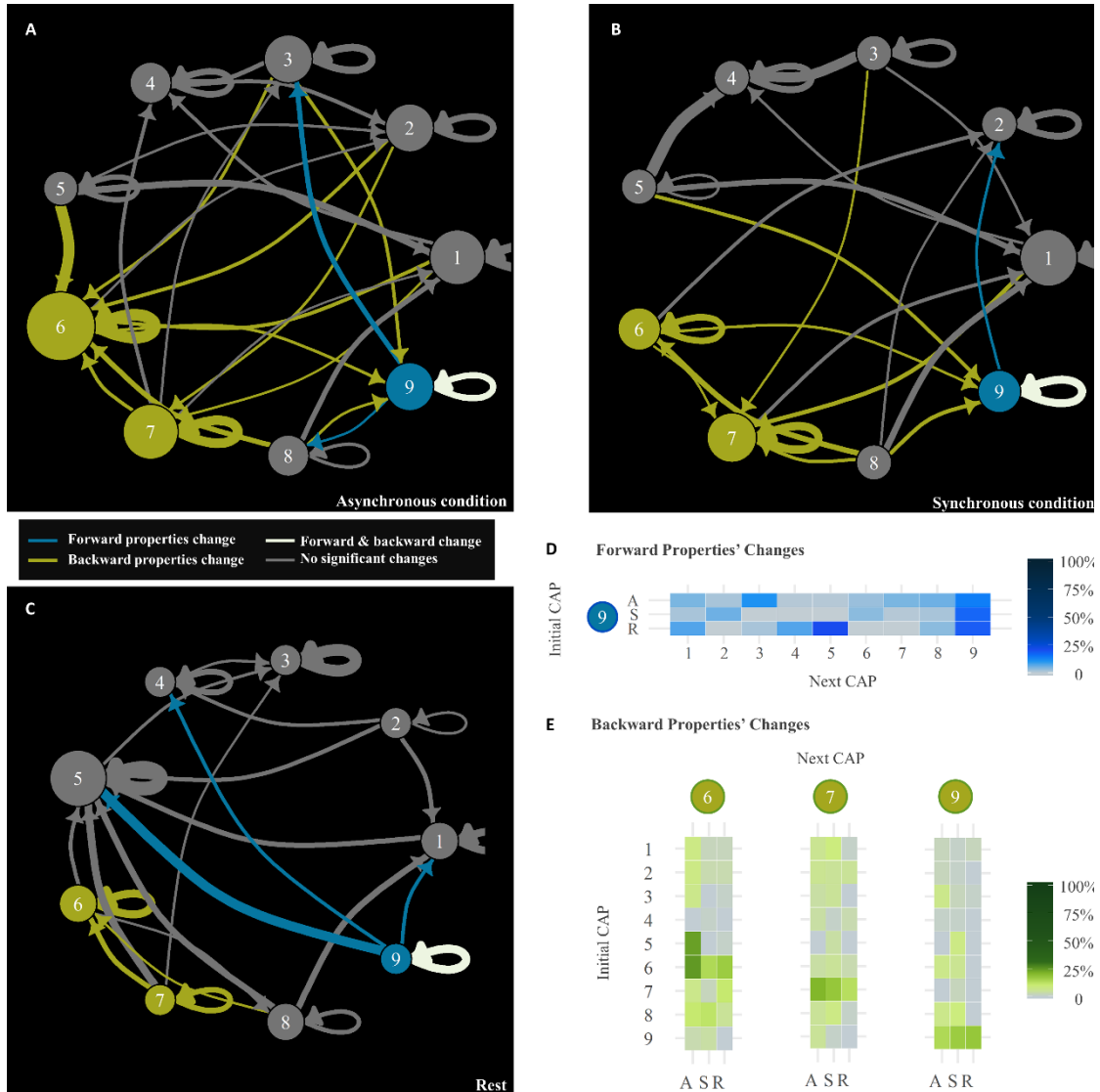


Figure 2.4 CAPs' transition dynamics across different conditions

Transitions that occur at least 5% of the time for each CAP, are shown across the different conditions. Highlighted in blue are CAPs that changed forward properties significantly across conditions. Transitions departing from such CAPs are also depicted in blue. Highlighted in green are CAPs that changed their backward properties significantly across conditions. Transitions arriving at such CAPs are also depicted in blue. Coloured in white, are the transitions departing from an initial CAP with significant changes in forward properties, and arriving at a next CAP with significant changes in backward properties. CAPs in grey did not change properties across conditions. The size of each CAP is proportional to the amount of arriving and departing transitions. The size of each transition is related to its probability. (A) Asynchronous condition. (B) Synchronous condition. (C) Rest. (D) changes in forward properties' across conditions (A - Asynchronous; S - Synchronous; R - Rest), can be seen in detail from CAPs 5 and 9 to all the other CAPs. (E) Changes in backward properties' across condition, can in turn be seen in detail for transitions from every CAP, to CAPs 2, 3, 5, 6, 7, and 9.

Part I: Identifying and modulating neural correlates of PH

significantly change the TPs, nor did its interaction with condition. Sensitivity to PE also did not show any significant effect ($F_{1, II}$ (PE)), nor did the interaction of this parameter with condition ($F_{1, III}$ (PE)). Overall, the experimental conditions, have a small effect on how the TPs change when departing from initial CAP 5, and from initial CAP 9, however no effect was observed for sensitivity to PH or to PE.

Backward properties. Here we focused on analysing the same parameters mentioned above, but for each fixed next CAP level ($CAP_{next} = CAP_i$). Models B_i I revealed a main effect of condition (figure 2.4A, 2.4B, 2.4C, Supplementary Figure S3) for transitions ending in CAP_{next} 2, CAP_{next} 5, CAP_{next} 6, and CAP_{next} 7 (all p-values < 0.001), as well as CAP_{next} 3 (p-value = 0.006), and CAP_{next} 9 (p-value = 0.006), as seen in figure 2.4E. PH sensitivity did not have a significant effect for any of the B_i II (PH) Models. However, Models B_i III (PH) detected a significant interaction effect between condition and PH sensitivity, for the transitions to CAP_{next} 2 (p-value = 0.012), and to CAP_{next} 6 (p-value = 0.012). Regarding PE, no main effect was found, but, crucially, a significant interaction between PE and condition was observed for the transitions to CAP_{next} 9 (B_i III (PE); p-value = 0.009). Overall, the experimental conditions, have a widespread effect on how the overall TPs of every CAP changes when transitioning to CAPs 2, 3, 5, 6, 7, and 9. Moreover, CAP 2 and CAP 6 were linked to PH and CAP 9 to PE.

Observing an interaction between PH sensitivity and condition implied that the effect of being sensitive to PH induction varied with the experimental condition, but only for TPs ending in the two above mentioned CAPs: CAP 2 and CAP 6. Hence, to better investigate the effect of PH sensitivity, we analysed each condition independently, for the transitions ending in CAP 2, and in CAP 6. Models B_i^c I (PH) revealed a significant effect PH sensitivity in the asynchronous condition for transitions ending in CAP_{next} 6 (estimate = 0.072, SE = ± 0.029 , p-value = 0.013, figure 2.5A, Supplementary Figure S4), but not in the synchronous condition (estimate = 0.008, SE = ± 0.016 , p-value = 0.59, figure 2.5B, Supplementary Figure S4), nor in rest (estimate = 0.012, SE = ± 0.018 , p-value = 0.47). Regarding transitions ending in CAP 2, no significant effect of sensitivity to PH induction was observed for any of the conditions. This showed that during PH induction, most CAPs will have an increase in the transition probability to CAP 6 (figure 2.3A). Consequently, and observed only in the asynchronous condition, the PH induction puts the brain's transition probabilities between different brain patterns, in a temporary dynamic arrangement, in which CAPs predominantly transition to CAP 6.

Observing an interaction between PE sensitivity and condition implied that the effect of being sensitive to PE induction varied with the experimental condition, only for the TPs ending in CAP 9. In the same way as for PH, we analysed the PE factor independently for each condition, for the transitions ending in CAP 9. Models B_9^c I (PE) revealed a significant effect of PE sensitivity on the transitions to CAP_{next} 9 in the asynchronous condition (estimate = -0.05, SE = ± 0.02 , p-value = 0.010, figure 2.5C, Supplementary Figure S5), but not in the synchronous condition (estimate = 0.005, SE = ± 0.017 , p-value = 0.76, figure 2.5D, Supplementary Figure S5) nor in rest (estimate = 0.002, SE = ± 0.008 , p-value = 0.77). This showed that if a participant is sensitive to PE induction, most CAPs will decrease their probability of transitioning to CAP 9, in the asynchronous condition.

Part I: Identifying and modulating neural correlates of PH

Brain activation for Co-Activation Pattern 9. The network identified as CAP 9 (figure 2.3B), contains significant clusters over the mPFC, precuneus, posterior cingulate cortex (PCC), bilateral AG, bilateral superior and middle frontal gyrus, as well as two bilateral clusters in the STS region. Subcortical clusters included bilateral thalamus and the bilateral cerebellum (Crus I and II).

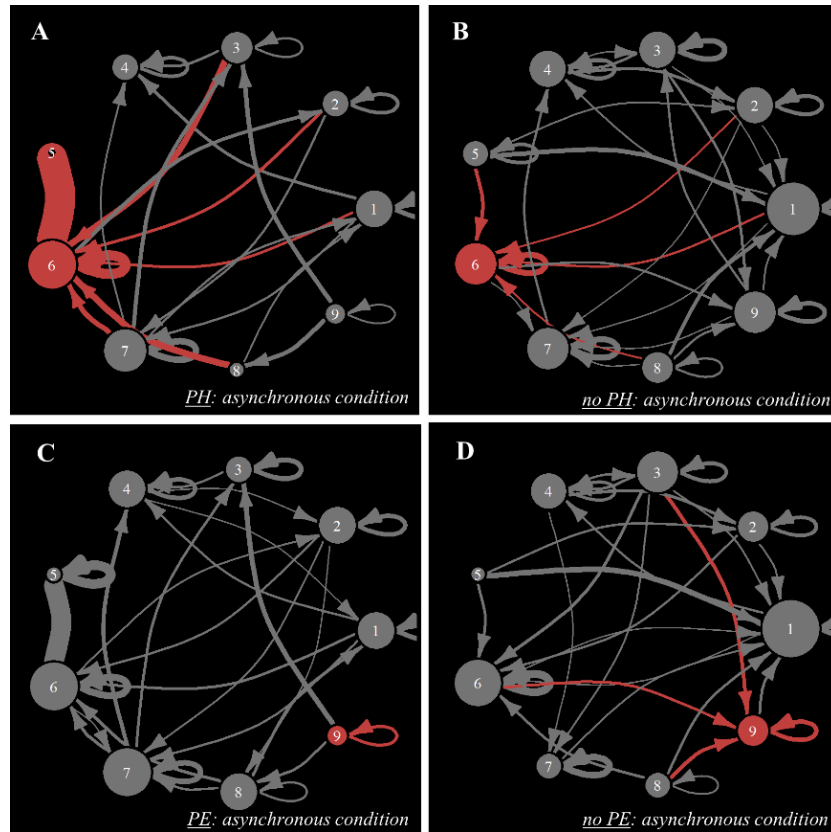


Figure 2.5 PH and PE transitions dynamics

Transitions that occur at least 5% of the time for each CAP, are shown in the asynchronous condition for different sensitivities to the induction of PH/PE. Highlighted in red are CAPs that significantly changed backward properties depending on the sensitivity to PH/PE. Transitions arriving at such CAPs are also depicted in red. **(A)** CAPs' transitions in the asynchronous condition when a PH is experienced **(B)** CAPs' transitions in the asynchronous condition without PH. The induction of a PH significantly increases the transition probabilities of the CAPs to CAP 6, resulting in a high convergence of the other CAPs to CAP 6, which is not seen if a PH is not experienced. **(C)** CAPs' transitions in the asynchronous condition when a PE is experienced **(D)** CAPs' transitions in the asynchronous condition without PE. Transitions arriving at such CAPs are also depicted in red. The induction of PE significantly decreases the transition probabilities of the CAPs to CAP 9.

2.1.5 Discussion

In the present work we used time-resolved analyses of fMRI brain activity to explore fluctuating brain states during a novel paradigm linking MR-compatible robotics and fMRI, used to control the subjective mental states of PH and PE in healthy individuals. We were able to identify two brain patterns, CAP 6 and CAP 9, which were both induced by the sensorimotor tasks, and sensitive to PH or PE, respectively. Such sensitivity was revealed as a temporal rearrangement of brain activity by altered transition probabilities that favoured CAP 6 in the event of PH and avoided CAP 9 in the event of PE. These data demonstrate that changes in transition dynamics of a specific network underlie the experience of a PH, and that such changes are independent of the experience of PE which in many participants accompanies PH.

2.1.5.1 Neural Correlates of PH-induction

A previous study has identified brain regions that in healthy participants are more active in the asynchronous condition (where PH and PE occur) rather than in the synchronous one, and further linked them with the PH through lesion network mapping analysis in neurological patients reporting PH (Bernasconi and Blondiaux et al., in press). Here, we identify a brain pattern, CAP 6, that has significantly higher occupancy and average duration in the asynchronous condition. We argue that this reveals the sensitivity of CAP 6 to the temporal delay present in the asynchronous condition, whereas other CAPs, such as 7 and 9, are modulated by the sensorimotor stimulation per se, and independently of the delay, given their higher occupancies and longer average durations for both asynchronous and synchronous conditions as compared to rest. These results show that asynchronous sensorimotor stimulation modulates a brain network that includes brain regions previously associated to PH (Bernasconi and Blondiaux et al., in press).

The findings presented here, specifically the analysis of the CAPs transition probabilities, further demonstrate that the induction of PH in the asynchronous condition is related to a significant change in how CAPs transition amongst themselves. Only during the asynchronous condition and only in the participants who are sensitive to PH-induction (i.e. higher PH ratings in the asynchronous versus synchronous condition) do we observe a significant increase in transition probabilities from all CAPs to CAP 6, as if this CAP is neurally “attracting” the other CAPs (Figure 5A). Accordingly, we argue that PH induction can be characterized by a perturbation of the “normal” brain network dynamics, consisting in an increase of the probability of all CAPs transitioning to CAP 6. No such changes in transition probabilities are observed in the synchronous condition nor in rest. Moreover, the fact that this CAP 6 change is not associated with the closely related conscious experience of PE, provides a disambiguation of the neural underpinnings of PH from those of PE (which are discussed in a subsequent section), even if PH is behaviourally, as tested by the present robotic system, typically accompanied by PE.

CAP 6 mainly consists of the right pSTS, the right dlPFC (including the IFG seed), the mPFC (including the SMA), and the bilateral AG. The former two areas (used as seeds in this analysis) have been previously linked to the PH (Arzy et al., 2006; Blanke et al., 2014; Bernasconi and Blondiaux et al., in press). In addition, the pSTS is a multisensory and sensorimotor brain region which responds to tactile, visual and auditory stimuli (Beauchamp et al., 2008), as well as

movement-related processes (Zito et al., 2020) and multisensory hallucinations (Ghazanfar and Schroeder, 2006). The mPFC cluster in CAP 6 has often been associated with different self-related processes (Gusnard et al., 2001; Platek et al., 2006; Beer et al., 2010; Whitfield-Gabrieli et al., 2011). In particular, a meta-analysis investigating the role of self-related, familiar, and other-related stimuli, identified this mPFC cluster as the main brain region distinguishing between self and other related processing (Qin and Northoff, 2011). We further note that the mPFC cluster includes the SMA, a region which has been implicated in bodily self-consciousness (i.e. Ferri et al., 2012; Ionta et al., 2013). The involvement of the bilateral AG has also been previously linked to PH, given that lesions in this area are associated with PH (Blanke et al., 2014). Finally, the deactivations we observe in the present study over the cuneus and occipital gyrus, which represent decreased BOLD in the secondary visual network (Shirer et al., 2012), are likely to be associated with opposing fluctuations between different functional networks (Fox and Raichle, 2007).

Disturbed interactions between functional networks have been hypothesized to represent a neural mechanism associated with several psychotic processes (Menon, 2011). Previous work suggested that hallucinations occur due to erroneous switches between two intrinsic networks, the central executive network (CEN) and default-mode network (DMN) (Menon and Uddin, 2010; Goulden et al., 2014). In particular, recent research showed that, when comparing hallucination and no-hallucination periods in schizophrenia, the interactions between major intrinsic networks follow different transition rules (Lefebvre et al., 2016). Similar mechanisms have also been proposed for Parkinson patients suffering from psychosis (Shine et al., 2014, 2015; Ffytche et al., 2017), a population typically afflicted by PH (Fénelon et al., 2011). The present data are compatible with that proposal. However, while previous work compared patients with hallucinations versus patients without hallucinations or patients with versus without psychosis, we report data in healthy participants in whom a specific hallucination, PH, is induced experimentally and in controlled fashion. Our data demonstrates, under the form of altered transition probabilities, that the induction of PH leads to aberrantly increased transitions to CAP 6, and that, by consequence, significantly different network interactions are observed between participants sensitive and those insensitive to PH induction. These results provide further evidence that hallucinations result from erroneous network switches (Goulden et al., 2014; Lefebvre et al., 2016) and extend this proposal to experimentally controlled hallucinations in healthy individuals.

2.1.5.2 Neural correlates of PE-induction

Most participants that were sensitivity to the induction of PH during the asynchronous condition of the sensorimotor task, experienced it with the accompanying sensation of PE. However, several participants that did not experience robot-induced PH in the asynchronous condition, nevertheless reported PE. Despite the strong link between these two sensations when being robotically induced in healthy individuals, the occurrence of PE without PH, paired with the use of dFC methods, allowed us to identify the different neural mechanisms that underlie both experiences. Although we observe three CAPs that have their occupancy and average duration modulated by the sensorimotor task (CAP 6, CAP 7, CAP 9), we find that only the change in the transition probabilities to CAP 9 is associated with PE. Distinctly from the induction of the PH, where all CAPs increase their transition probabilities to CAP 6, the induction of PE in the asynchronous condition is characterised by a significant decrease of the

transition probabilities of all CAPs to CAP 9, suggesting that this brain pattern is generally avoided in the event of PE. No such significant changes in transition probabilities to CAP 9 are observed in the synchronous condition nor in rest.

CAP 9 overlaps in three main regions with CAP 6; i.e., the pSTS, mPFC, and bilateral AG. Besides the implications that have been proposed for the latter component in PH, we consider a possible dual role of this region, given that the AG has been extensively implicated in PE both in healthy individuals (Farrer and Frith, 2002; Blakemore et al., 2003), and in patients with schizophrenia and symptomatic passivity experiences (Farrer et al., 2004). In addition to these common regions, CAP 9 also includes the precuneus, the PCC, extensions of the pSTS activation over the middle temporal gyrus, and the mPFC cluster, which here does not include the SMA and is significantly larger than in CAP 6, extending to midline cortical structures over the ventral and dorsal medial prefrontal cortices. The presence of the bilateral AG, precuneus, PPC, together with the observed midline cortical structures, which overlap significantly with the midline cortical structures of the DMN (Raichle et al., 2001; Shirer et al., 2012), suggest that the brain pattern of CAP 9 is closely related to the DMN. This is consistent with previous studies using the same or similar dFC methods, which recovered either parts or the complete DMN (Kiviniemi et al., 2011; Liu and Duyn, 2013; Liu et al., 2018). Diminished network interactions with the DMN in hallucinations have been reported before; e.g., in first-episode psychosis patients (Jardri et al., 2013), and in schizophrenic patients with positive symptoms (Lefebvre et al., 2016), both spontaneously hallucinating during resting state fMRI. The effect observed here for PE, with most CAPs decreasing their transition probabilities to CAP 9, contrasts with our initial prediction that PE, as PH, would be grounded on a dominance of a brain state over another. However, shunning of specific networks is known in hallucinations, particularly for the DMN, as observed in psychotic patients (Jardri et al., 2013; Lefebvre et al., 2016). Due to the very significant number of features shared between the identified CAP 9 for which this mechanism occurs associated with PE, and the DMN, we do not exclude the possibility that what is detected here for PE-induction might be in fact a more general mechanism underlying hallucinations, beyond PE.

2.1.5.3 Mechanisms of PH and PE in robotically mediated induction of hallucinations

In the present work, we find that both the induction of PH and PE are characterized by significant changes in network interactions. We previously described two regions, pSTS and IFG, that were more active during robot-induced PH in the asynchronous condition (Bernasconi and Blondiaux et al., in press) and that also overlapped with an impaired functional network as defined in neurological patients with clinically-relevant PH. Based on the present findings, we propose that the activations of the pSTS and IFG in the asynchronous condition represent a general predisposition to the variations in network behaviour observed for CAP 6 and CAP 9, which lead to PH and PE respectively. Hence, in the asynchronous condition, the predisposition to having PH or PE is marked by such activations of the pSTS and IFG, but, importantly, PH and PE will only occur once the changes in network behaviour also occur. This view is supported by observations that certain brain regions can have a prominent role in the switching between different networks (Sridharan et al., 2008; Manoliu et al., 2014). In addition, hallucination processes in psychopathology also lend support to this hypothesis. Positive symptoms in schizophrenia and schizotypal disorders, can be marked by constant

dysfunctions in functional connectivity between brain regions, that correlate with hallucination severity (Fletcher and Frith, 2009; Skudlarski et al., 2010; Ettinger et al., 2015), however, the occurrence of hallucinations is limited to time periods characterized by changes in network interactions (Lefebvre et al., 2016).

2.1.5.4 Methodological considerations

The study of dynamics of brain activity as measured by fMRI has received considerable attention during the past decade (Hutchison et al., 2013; Preti et al., 2017) and is particularly relevant to explain complex behaviour and psychopathology (Bolton et al., 2020a). In this work, we opted for CAPs analysis, which starts from the selection of relevant seeds, to probe their interaction with the rest of the brain in terms of dynamically occurring co-activation patterns. The most interesting alterations found in this study were characterized by transition probabilities, a rather subtle correlate of brain activity that has not yet been exploited to a big extent. More work is thus needed on this topic. There are also a number of limitations. First, we focused our analysis by choosing as seeds, two key brain regions previously identified for PH. While this approach allowed us to narrow down the scope of the measures, it might also be that other processes in the brain were missed. Second, temporal sequence analysis could be extended by generative models (Bolton et al., 2018; Vidaurre et al., 2017; Zhang et al., 2020), which can even be applied to individual nodes instead of at the network level (Bolton et al., 2020c). Finally, another option would be to apply effective connectivity models, such as dynamic causal modelling (Friston, 2011) or Granger causality (Valdés-Sosa et al., 2005), which are used to infer inter-regional interactions supported by anatomical connections.

2.1.5.5 Conclusions

In sum, we identify dynamic fluctuations of brain activity that underlie PH and PE in healthy participants. We show that, the robot-induced sensations in the asynchronous condition are characterized by subtle changes in brain pattern transitions. Whereas the asynchronous condition is characterized by increased activations of the pSTS and IFG, as well as higher occupancy and average duration of the network CAP 6, for PH, we identify a significant increase of the probabilities for all observed brain patterns to transition to CAP 6, and for PE, we observe an avoidance of all observed brain patterns to transition to a partly overlapping, but different network, CAP 9. These results highlight the subtle neural changes of a specific network during robot-induced PH in healthy individuals and further disambiguate the brain processes of PH from those of the typically accompanying PE. Furthermore, we extend changes in network behaviour associated with clinically relevant hallucinations (Menon, 2011; Jardri et al., 2013; Lefebvre et al., 2016) to network behaviour during an experimentally-controlled specific hallucination, PH.

Author Contributions

Herberto Dhanis: Conceptualization, Methodology, Software, Formal analysis, Data curation, Writing - original draft and editing, Visualization. **Thomas Bolton:** Methodology, Supervision. **Giulio Rognini:** Experimental methodology. **Eva Blondiaux and Nathan Faivre:** Supervision. **Dimitri Van De Ville and Olaf Blanke:** Conceptualization, Supervision, Writing – Review & Editing.

Declarations of interest: none.

Acknowledgments

This work presented here was supported in by the Bertarelli Foundation, the Swiss National Science Foundation, the National Centres of Competence in Research, as well as by two generous donors advised by Carigest SA. The first of such donors wishes to remain anonymous, whilst the second one is the Fondazione Teofilo Rossi di Montelera e di Premuda.

2.1.6 References

- Arzy, S., Seeck, M., Ortigue, S., Spinelli, L., Blanke, O., 2006a. Induction of an illusory shadow person. *Nature* 443, 287. <https://doi.org/10.1038/443287a>
- Arzy, S., Thut, G., Mohr, C., Michel, C.M., Blanke, O., 2006b. Neural basis of embodiment: Distinct contributions of temporoparietal junction and extrastriate body area. *J. Neurosci.* 26, 8074–8081. <https://doi.org/10.1523/JNEUROSCI.0745-06.2006>
- Bates, D., Mächler, M., Bolker, B.M., Walker, S.C., 2015. Fitting linear mixed-effects models using lme4. *J. Stat. Softw.* 67. <https://doi.org/10.18637/jss.v067.i01>
- Beauchamp, M.S., Yasar, N.E., Frye, R.E., Ro, T., 2008. Touch, sound and vision in human superior temporal sulcus. *Neuroimage* 41, 1011–1020. <https://doi.org/10.1016/j.neuroimage.2008.03.015>
- Beer, J.S., Lombardo, M. V., Bhanji, J.P., 2010. Roles of medial prefrontal cortex and orbitofrontal cortex in self-evaluation. *J. Cogn. Neurosci.* 22, 2108–2119. <https://doi.org/10.1162/jocn.2009.21359>
- Benjamini, Y., Hochberg, Y., 1995. Controlling the false discovery rate: A practical and powerful approach to multiple testing. *R. Stat. Soc.* 57, 289–300.
- Bernasconi, F., Blondiaux, E., Potheegadoo, J., Stripeikyte, G., 2020a. Sensorimotor hallucinations in Parkinson's disease. *bioRxiv* 1–38. <https://doi.org/https://doi.org/10.1101/2020.05.11.054619>
- Bernasconi, F., Blondiaux, E., Potheegadoo, J., Stripeikyte, G., 2020b. Sensorimotor hallucinations in Parkinson's disease. *bioRxiv*.
- Blakemore, S.J., Oakley, D.A., Frith, C.D., 2003. Delusions of alien control in the normal brain. *Neuropsychologia* 41, 1058–1067. [https://doi.org/10.1016/S0028-3932\(02\)00313-5](https://doi.org/10.1016/S0028-3932(02)00313-5)
- Blanke, O., 2012. Multisensory brain mechanisms of bodily self-consciousness. *Nat. Rev. Neurosci.* 13, 556–571. <https://doi.org/10.1038/nrn3292>
- Blanke, O., Arzy, S., Landis, T., 2008. Chapter 22 Illusory reduplications of the human body and self. *Handb. Clin. Neurol.* 88, 429–458. [https://doi.org/10.1016/S0072-9752\(07\)88022-5](https://doi.org/10.1016/S0072-9752(07)88022-5)
- Blanke, O., Ortigue, S., Coeytaux, A., Martory, M.D., Landis, T., 2003. Hearing of a presence. *Neurocase* 9, 329–339. <https://doi.org/10.1076/neur.9.4.329.15552>
- Blanke, O., Pozeg, P., Hara, M., Heydrich, L., Serino, A., Yamamoto, A., Higuchi, T., Salomon, R., Seeck, M., Landis, T., Arzy, S., Herbelin, B., Bleuler, H., Rognini, G., 2014a. Neurological and Robot-Controlled Induction of an Apparition. *Curr. Biol.* 24, 2681–2686. <https://doi.org/10.1016/j.cub.2014.09.049>
- Blanke, O., Pozeg, P., Hara, M., Heydrich, L., Serino, A., Yamamoto, A., Higuchi, T., Salomon, R., Seeck, M., Landis, T., Arzy, S., Herbelin, B., Bleuler, H., Rognini, G., 2014b. Neurological and robot-controlled induction of an apparition. *Curr. Biol.* 24, 2681–2686. <https://doi.org/10.1016/j.cub.2014.09.049>
- Bolton, T.A.W., Morgenroth, E., Preti, M.G., Van De Ville, D., 2020a. Tapping into Multi-Faceted Human Behavior and Psychopathology Using fMRI Brain Dynamics. *Trends Neurosci.* 43, 667–680. <https://doi.org/10.1016/j.tins.2020.06.005>

Part I: Identifying and modulating neural correlates of PH

- Bolton, T.A.W., Tarun, A., Sterpenich, V., Schwartz, S., Van De Ville, D., 2018. Interactions between Large-Scale Functional Brain Networks are Captured by Sparse Coupled HMMs. *IEEE Trans. Med. Imaging* 37, 230–240. <https://doi.org/10.1109/TMI.2017.2755369>
- Bolton, T.A.W., Tuleasca, C., Rey, G., Wotruba, D., Gaviria, J., Dhanis, H., Blondiaux, E., Gauthier, B., Smigielski, L., Van De Ville, D., 2019. TbCAPs: A ToolBox for Co-Activation Pattern Analysis.
- Bolton, T.A.W., Tuleasca, C., Wotruba, D., Rey, G., Dhanis, H., Gauthier, B., Delavari, F., Morgenroth, E., Gaviria, J., Blondiaux, E., Smigielski, L., Van De Ville, D., 2020b. TbCAPs: A toolbox for co-activation pattern analysis. *Neuroimage* 211, 116621. <https://doi.org/10.1016/j.neuroimage.2020.116621>
- Bolton, T.A.W., Urunuela, E., Tian, Y., Zalesky, A., Caballero-Gaudes, C., Van De Ville, D., 2020c. Sparse coupled logistic regression to estimate co-activation and modulatory influences of brain regions. *J. Neural Eng.* 17. <https://doi.org/10.1088/1741-2552/aba55e>
- Brugger, P., Regard, M., Landis, T., 1997. Illusory reduplication of one's own body: Phenomenology and classification of autoscopic phenomena. *Cogn. Neuropsychiatry* 2, 19–38. <https://doi.org/10.1080/135468097396397>
- Brugger, P., Regard, M., Landis, T., 1996. Unilaterally felt “presences”: The neuropsychiatry of one's invisible doppelgänger. *Neuropsychiatry, Neuropsychol. Behav. Neurol.* 9, 114–122.
- Chen, J.E., Chang, C., Greicius, M.D., Glover, G.H., 2015. Introducing co-activation pattern metrics to quantify spontaneous brain network dynamics. *Neuroimage* 111, 476–488. <https://doi.org/10.1016/j.neuroimage.2015.01.057>
- Damaraju, E., Allen, E.A., Belger, A., Ford, J.M., McEwen, S., Mathalon, D.H., Mueller, B.A., Pearlson, G.D., Potkin, S.G., Preda, A., Turner, J.A., Vaidya, J.G., Van Erp, T.G., Calhoun, V.D., 2014. Dynamic functional connectivity analysis reveals transient states of dysconnectivity in schizophrenia. *NeuroImage Clin.* 5, 298–308. <https://doi.org/10.1016/j.nicl.2014.07.003>
- Diederich, N.J., Fénelon, G., Stebbins, G., Goetz, C.G., 2009. Hallucinations in Parkinson disease. *Nat. Rev. Neurol.* 5, 331–342. <https://doi.org/10.1038/nrneurol.2009.62>
- Ettinger, U., Mohr, C., Gooding, D.C., Cohen, A.S., Rapp, A., Haenschel, C., Park, S., 2015. Cognition and brain function in schizotypy: A selective review. *Schizophr. Bull.* 41, S417–S426. <https://doi.org/10.1093/schbul/sbu190>
- Farrer, C., Franck, N., Frith, C.D., Decety, J., Georgieff, N., D'Amato, T., Jeannerod, M., 2004. Neural correlates of action attribution in schizophrenia. *Psychiatry Res. - Neuroimaging* 131, 31–44. <https://doi.org/10.1016/j.psychresns.2004.02.004>
- Farrer, C., Frith, C.D., 2002. Experiencing oneself vs another person as being the cause of an action: The neural correlates of the experience of agency. *Neuroimage* 15, 596–603. <https://doi.org/10.1006/nimg.2001.1009>
- Fénelon, G., Soulas, T., De Langavant, L.C., Trinkler, I., Bachoud-Lévi, A.C., 2011. Feeling of presence in Parkinson's disease. *J. Neurol. Neurosurg. Psychiatry* 82, 1219–1224. <https://doi.org/10.1136/jnnp.2010.234799>

Part I: Identifying and modulating neural correlates of PH

- Ferri, F., Frassinetti, F., Ardizzi, M., Costantini, M., Gallese, V., 2012. A sensorimotor network for the bodily self. *J. Cogn. Neurosci.* 24, 1584–1595. https://doi.org/10.1162/jocn_a_00230
- Ffytche H, D., Creese, B., Politis, M., Chaudhuri, K.R., Weintraub, D., Ballard, C., Aarsland, D., 2017. The psychosis spectrum in Parkinson disease. *Nat. Rev. Neurol.* <https://doi.org/10.1038/nrneurol.2016.200>
- Fletcher, P.C., Frith, C.D., 2009. Perceiving is believing: A Bayesian approach to explaining the positive symptoms of schizophrenia. *Nat. Rev. Neurosci.* 10, 48–58. <https://doi.org/10.1038/nrn2536>
- Fox, M.D., Raichle, M.E., 2007. Spontaneous fluctuations in brain activity observed with functional magnetic resonance imaging. *Nat. Rev. Neurosci.* 8, 700–711. <https://doi.org/10.1038/nrn2201>
- Friston, K.J., 2011. Functional and Effective Connectivity: A Review. *Brain Connect.* 1, 13–36. <https://doi.org/10.1089/brain.2011.0008>
- Ghazanfar, A.A., Schroeder, C.E., 2006. Is neocortex essentially multisensory? *Trends Cogn. Sci.* 10, 278–285. <https://doi.org/10.1016/j.tics.2006.04.008>
- Goulden, N., Khusnulina, A., Davis, N.J., Bracewell, R.M., Bokde, A.L., McNulty, J.P., Mullins, P.G., 2014. The salience network is responsible for switching between the default mode network and the central executive network: Replication from DCM. *Neuroimage* 99, 180–190. <https://doi.org/10.1016/j.neuroimage.2014.05.052>
- Gusnard, D.A., Akbudak, E., Shulman, G.L., Raichle, M.E., 2001. Medial prefrontal cortex and self-referential mental activity: Relation to a default mode of brain function. *Proc. Natl. Acad. Sci. U. S. A.* 98, 4259–4264. <https://doi.org/10.1073/pnas.071043098>
- Hutchison, R.M., Womelsdorf, T., Allen, E.A., Bandettini, P.A., Calhoun, V.D., Corbetta, M., Della Penna, S., Duyn, J.H., Glover, G.H., Gonzalez-Castillo, J., Handwerker, D.A., Keilholz, S., Kiviniemi, V., Leopold, D.A., de Pasquale, F., Sporns, O., Walter, M., Chang, C., 2013. Dynamic functional connectivity: Promise, issues, and interpretations. *Neuroimage* 80, 360–378. <https://doi.org/10.1016/j.neuroimage.2013.05.079>
- Ionta, S., Martuzzi, R., Salomon, R., Blanke, O., 2013. The brain network reflecting bodily self-consciousness: A functional connectivity study. *Soc. Cogn. Affect. Neurosci.* 9, 1904–1913. <https://doi.org/10.1093/scan/nst185>
- James, W., 1902. *A study of Man: The varieties of religious experiences*. Longmans, Green Co. <https://doi.org/10.1353/jsp.2003.0021>
- Jardri, R., Thomas, P., Delmaire, C., Delion, P., Pins, D., 2013. The neurodynamic organization of modality-dependent hallucinations. *Cereb. Cortex* 23, 1108–1117. <https://doi.org/10.1093/cercor/bhs082>
- Jaspers, K., 1913. Über leibhafte Bewusstheiten (Bewusstheitstäuschungen), ein psychopathologisches Elementarsymptom. *Z Pathopsychol* 2, 150–161.

Part I: Identifying and modulating neural correlates of PH

- Kiviniemi, V., Vire, T., Remes, J., Elseoud, A.A., Starck, T., Tervonen, O., Nikkinen, J., 2011. A Sliding Time-Window ICA Reveals Spatial Variability of the Default Mode Network in Time. *Brain Connect.* 1, 339–347. <https://doi.org/10.1089/brain.2011.0036>
- Lefebvre, S., Demeulemeester, M., Leroy, A., Delmaire, C., Lopes, R., Pins, D., Thomas, P., Jardri, R., 2016. Network dynamics during the different stages of hallucinations in schizophrenia. *Hum. Brain Mapp.* 37, 2571–2586. <https://doi.org/10.1002/hbm.23197>
- Liu, X., 2016. Chapter 3 - Linear mixed-effects models, in: Liu, X.B.T.-M. and A. of L.D.A. (Ed.), . Academic Press, Oxford, pp. 61–94. <https://doi.org/https://doi.org/10.1016/B978-0-12-801342-7.00003-4>
- Liu, X., Duyn, J.H., 2013. Time-varying functional network information extracted from brief instances of spontaneous brain activity. *Proc. Natl. Acad. Sci. U. S. A.* 110, 4392–4397. <https://doi.org/10.1073/pnas.1216856110>
- Liu, X., Zhang, N., Chang, C., Duyn, J.H., 2018. NeuroImage Co-activation patterns in resting-state fMRI signals. *Neuroimage* 180, 485–494. <https://doi.org/10.1016/j.neuroimage.2018.01.041>
- Manoliu, A., Riedl, V., Zherdin, A., Mühlau, M., Schwerthöffer, D., Scherr, M., Peters, H., Zimmer, C., Förstl, H., Bäuml, J., Wohlschläger, A.M., Sorg, C., 2014. Aberrant dependence of default mode/central executive network interactions on anterior insular salience network activity in schizophrenia. *Schizophr. Bull.* 40, 428–437. <https://doi.org/10.1093/schbul/sbt037>
- Matsushashi, M., Ikeda, A., Ohara, S., Matsumoto, R., Yamamoto, J., Takayama, M., Satow, T., Begum, T., Usui, K., Nagamine, T., Mikuni, N., Takahashi, J., Miyamoto, S., Fukuyama, H., Shibasaki, H., 2004. Multisensory convergence at human temporo-parietal junction - Epicortical recording of evoked responses. *Clin. Neurophysiol.* 115, 1145–1160. <https://doi.org/10.1016/j.clinph.2003.12.009>
- Menon, V., 2011. Large-scale brain networks and psychopathology: A unifying triple network model. *Trends Cogn. Sci.* 15, 483–506. <https://doi.org/10.1016/j.tics.2011.08.003>
- Menon, V., Uddin, L.Q., 2010. Saliency, switching, attention and control: a network model of insula function. *Brain Struct. Funct.* 214, 655–667. <https://doi.org/10.1007/s00429-010-0262-0>
- Messner, R., 2016. *Naked Mountain: Nanga Parbat, Brother, Death, Solitude.* Crowood.
- Mlakar, J., Jensterle, J., Frith, C.D., 1994. Central monitoring deficiency. *Psychol. Med.* 24, 557–564.
- Monti, S., Tamayo, P., Mesirov, J., Golub, T., 2003. Consensus clustering: A resampling-based method for class discovery and visualization of gene expression microarray data. *Mach. Learn.* 52, 91–118. <https://doi.org/10.1023/A:1023949509487>
- Oldfield, R.C., 1971. The assessment and analysis of handedness: The Edinburgh Inventory. *Neuropsychologia* 9, 97–113. https://doi.org/10.1007/978-0-387-79948-3_6053
- Platek, S.M., Loughhead, J.W., Gur, R.C., Busch, S., Ruparel, K., Phend, N., Panyavin, I.S., Langleben, D.D., 2006. Neural substrates for functionally discriminating self-face from personally familiar faces. *Hum. Brain Mapp.* 27, 91–98. <https://doi.org/10.1002/hbm.20168>

Part I: Identifying and modulating neural correlates of PH

- Power, J.D., Barnes, K.A., Snyder, A.Z., Schlaggar, B.L., Petersen, S.E., 2012. Spurious but systematic correlations in functional connectivity MRI networks arise from subject motion. *Neuroimage* 59, 2142–2154. <https://doi.org/10.1016/j.neuroimage.2011.10.018>
- Preti, M.G., Bolton, T.A., Van De Ville, D., 2017. The dynamic functional connectome: State-of-the-art and perspectives. *Neuroimage* 160, 41–54. <https://doi.org/10.1016/j.neuroimage.2016.12.061>
- Qin, P., Northoff, G., 2011. How is our self related to midline regions and the default-mode network? *Neuroimage* 57, 1221–1233. <https://doi.org/10.1016/j.neuroimage.2011.05.028>
- Raichle, M.E., MacLeod, A.M., Snyder, A., Powers, W.J., Gusnard, D.A., 2001. A default mode of brain function. *Proc. Natl. Acad. Sci.* 98, 676–682.
- Rashid, B., Damaraju, E., Pearlson, G.D., Calhoun, V.D., 2014. Dynamic connectivity states estimated from resting fMRI Identify differences among Schizophrenia, bipolar disorder, and healthy control subjects. *Front. Hum. Neurosci.* 8, 1–13. <https://doi.org/10.3389/fnhum.2014.00897>
- Shine, J.M., Muller, A.J., O’Callaghan, C., Hornberger, M., Halliday, G.M., Lewis, S.J.G., 2015. Abnormal connectivity between the default mode and the visual system underlies the manifestation of visual hallucinations in Parkinson’s disease: A task-based fMRI study. *Parkinsons. Dis.* 1. <https://doi.org/10.1038/npjparkd.2015.3>
- Shine, J.M., O’Callaghan, C., Halliday, G.M., Lewis, S.J.G., 2014. Tricks of the mind: Visual hallucinations as disorders of attention. *Prog. Neurobiol.* 116, 58–65. <https://doi.org/10.1016/j.pneurobio.2014.01.004>
- Shirer, W R, Ryali, S., Rykhlevskaia, E., Menon, V., Greicius, M.D., 2012. Decoding Subject-Driven Cognitive States with Whole-Brain Connectivity Patterns 3, 158–165. <https://doi.org/10.1093/cercor/bhr099>
- Shirer, W. R., Ryali, S., Rykhlevskaia, E., Menon, V., Greicius, M.D., 2012. Decoding subject-driven cognitive states with whole-brain connectivity patterns. *Cereb. Cortex* 22, 158–165. <https://doi.org/10.1093/cercor/bhr099>
- Skudlarski, P., Jagannathan, K., Anderson, K., Stevens, M.C., Calhoun, V.D., Skudlarska, B.A., Pearlson, G., 2010. Brain Connectivity Is Not Only Lower but Different in Schizophrenia: A Combined Anatomical and Functional Approach. *Biol. Psychiatry* 68, 61–69. <https://doi.org/10.1016/j.biopsych.2010.03.035>
- Sridharan, D., Levitin, D.J., Menon, V., 2008. A critical role for the right fronto-insular cortex in switching between central-executive and default-mode networks. *Proc. Natl. Acad. Sci. U. S. A.* 105, 12569–12574. <https://doi.org/10.1073/pnas.0800005105>
- Tagliazucchi, E., Balenzuela, P., Fraiman, D., Chialvo, D.R., 2012. Criticality in large-scale brain fmri dynamics unveiled by a novel point process analysis. *Front. Physiol.* 3 FEB, 1–12. <https://doi.org/10.3389/fphys.2012.00015>
- Tagliazucchi, E., Balenzuela, P., Fraiman, D., Montoya, P., Chialvo, D.R., 2011. Spontaneous BOLD event triggered averages for estimating functional connectivity at resting state. *Neurosci. Lett.* 488, 158–163. <https://doi.org/10.1016/j.neulet.2010.11.020>

Part I: Identifying and modulating neural correlates of PH

Valdés-Sosa, P.A., Sánchez-Bornot, J.M., Lage-Castellanos, A., Vega-Hernández, M., Bosch-Bayard, J., Melie-García, L., Canales-Rodríguez, E., 2005. Estimating brain functional connectivity with sparse multivariate autoregression. *Philos. Trans. R. Soc. B Biol. Sci.* 360, 969–981. <https://doi.org/10.1098/rstb.2005.1654>

Vidaurre, D., Smith, S.M., Woolrich, M.W., 2017. Brain network dynamics are hierarchically organized in time 114. <https://doi.org/10.1073/pnas.1705120114>

Whitfield-Gabrieli, S., Moran, J.M., Nieto-Castañón, A., Triantafyllou, C., Saxe, R., Gabrieli, J.D.E., 2011. Associations and dissociations between default and self-reference networks in the human brain. *Neuroimage* 55, 225–232. <https://doi.org/10.1016/j.neuroimage.2010.11.048>

Zhang, G., Cai, B., Zhang, A., Stephen, J.M., Wilson, T.W., Calhoun, V.D., Wang, Y.P., 2020. Estimating Dynamic Functional Brain Connectivity with a Sparse Hidden Markov Model. *IEEE Trans. Med. Imaging* 39, 488–498. <https://doi.org/10.1109/TMI.2019.2929959>

Zito, G.A., Wiest, R., Aybek, S., 2020. Neural correlates of sense of agency in motor control: A neuroimaging meta-analysis. *PLoS One* 15, 1–17. <https://doi.org/10.1371/journal.pone.0234321>

2.1.7 Supplementary Information

Table 2.1 Questionnaire on robot-induced sensations

Question ID	Question Description	Illusory Perception Assessed
1	"I felt as if I was touching my body"	Self-touch
2	"I felt as if I was touching someone else's body"	Control question
3	"I felt as if I had no body"	Control question
4	"I felt as if I had two right hands"	Control question
5	"I felt as if someone else was touching my body"	Passivity Experience
6	"I felt as if someone was behind me"	Presence Hallucination

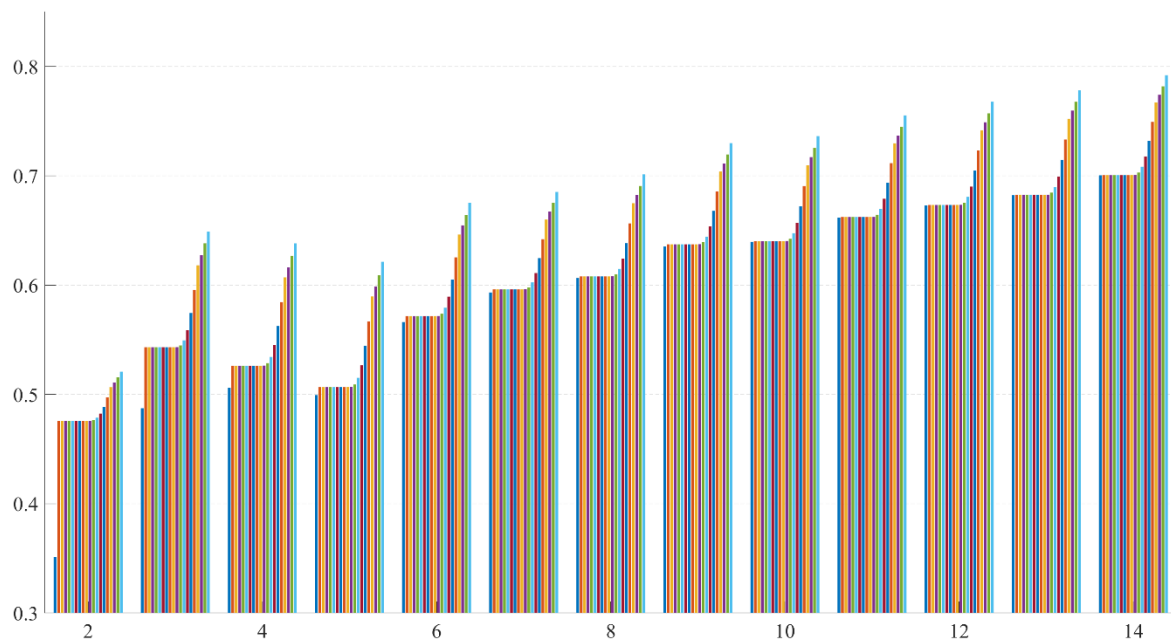


Figure 2.6 [S1] Consensus clustering

Output from the TbCaps (Bolton, et al. 2020) computed for a different number of centroids (k). On the y-axis, the stability measure is computed from the number of ambiguously clustered pairs (PAC), 1-PAC. Different colours represent different consensus thresholds, used when computing the PAC. The number of centroids was selected at 9, given that besides the general tendency of Consensus Clustering to improve with increasing k (Monti et al., 2003), it seemed to be the k that increased stability the most, at the same time as not seeing an increase for the k just after it.

Part I: Identifying and modulating neural correlates of PH

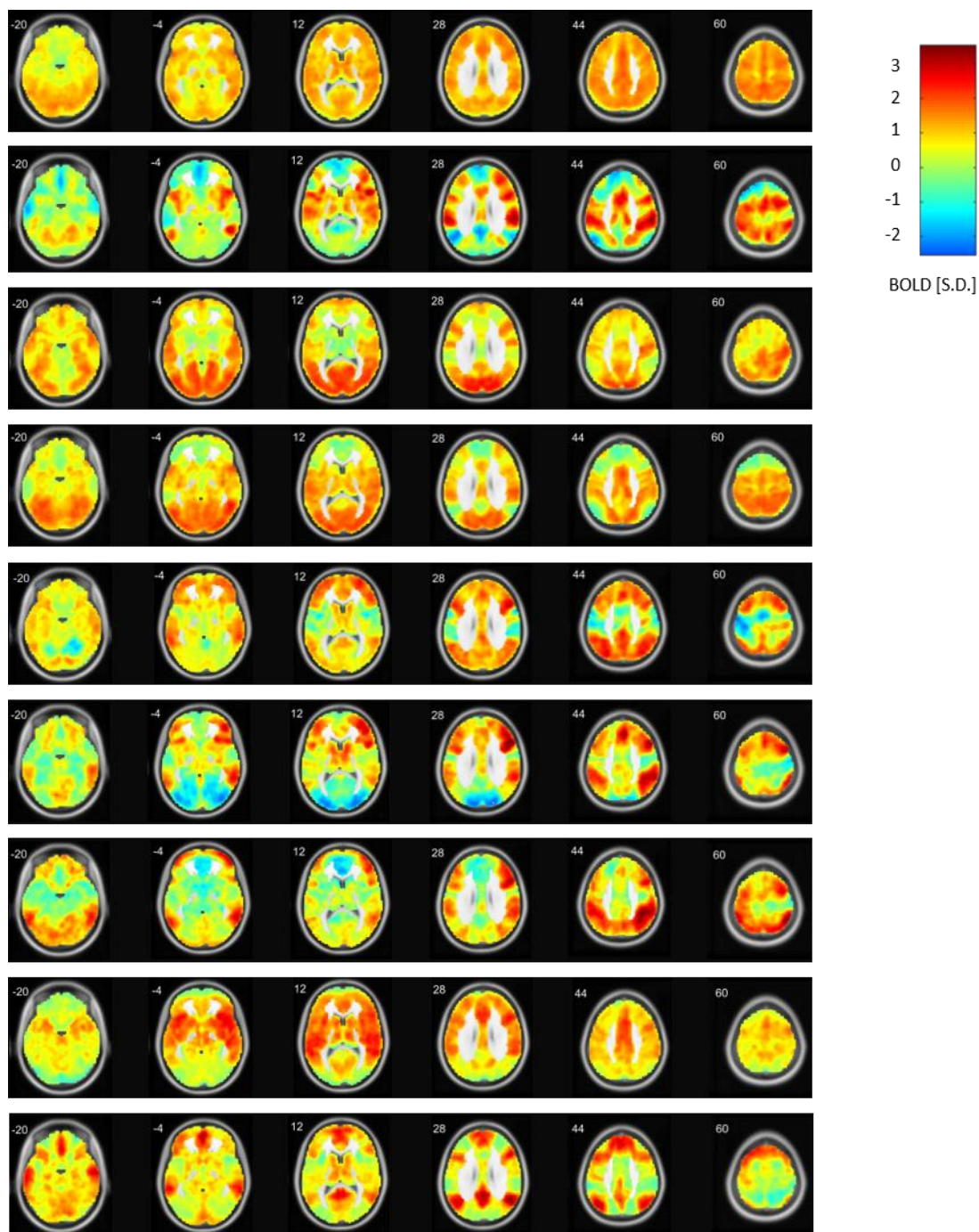


Figure 2.7 [S2] CAPs from 1(top) to 9 (bottom) after spatial z-scoring

Part I: Identifying and modulating neural correlates of PH

Table 2.2 [S2] CAPs' cluster peaks

Regions corresponding to the positive cluster peaks of all CAPs. Main regions were considered regions for which the spatial standard deviation, exceeded at least half of the maximum standard deviation, and which included at least 50 voxels. * CAP 1 is not described here as it is a mild widespread activation CAP which is typical to arise when using CAP analysis, but that does not yield any particular meaningful information (Bolton, T., personal communication).

CAP	Main Regions [$Z > \frac{z_{max}}{2}$ & > 50 Voxels]	Voxels	BA	Peak MNI coordinates		
				x	y	z
1	*					
2	R. Posterior Superior Temporal Sulcus (seed)	524		54	-52	0
	R. Inferior Frontal Gyrus (includes IFG seed)	12781	6 / 40 / 24 / 13 / 2 / 32	50	10	20
	L. Insula	1470	13 / 44 / 22	-32	16	8
	L. Middle Temporal Gyrus	158		-52	-62	0
	R. Middle Frontal Gyrus	1252	10 / 46 / 9	36	40	34
	R. Supramarginal Gyrus	4335	40 / 2 / 3	58	-34	36
	R. Middle Cingulate Gyrus	249	31	12	-26	40
	R. Precuneus	832	7 / 5	10	-60	58
3	Superior Temporal Gyrus; Middle Temporal Gyrus; Middle Occipital Gyrus; Lingual Gyrus; Precentral Gyrus; Postcentral Gyrus; Precuneus; Cuneus; Superior Parietal Lobule. Extends to: Inferior Frontal Gyrus and Medial Frontal Gyrus	47619	19 / 18 / 7 / 21 / 22 / 37 / 31 / 6 / 39 / 4	-14	-88	32
	L. Hippocampus	65		-26	-14	-22
	Medial Frontal Orbital Gyrus	1215	10 / 11	-2	66	-2
	L. Inferior Frontal Triangularis	718	9 / 45 / 46	-54	22	22
	R. Insula	197	13	38	-14	20
	L. Postcentral	610	4 / 6	-60	-6	22
4	Superior Temporal Gyrus; Middle Temporal Gyrus; Middle Occipital Gyrus; Lingual Gyrus; Precentral Gyrus; Postcentral Gyrus; Inferior Parietal Lobule; Precuneus; Posterior Lobe of the Cerebellum; Anterior Lobe of the Cerebellum; Cuneus; Superior Parietal Lobule. Extends to: Superior frontal Gyrus, Supplemental Motor Area, and Middle Frontal Gyrus	49336	19 / 18 / 7 / 21 / 22 / 37 / 31 / 6 / 39 / 4	54	-56	0
	L. Insula	165		-38	-18	-2
	R. Middle Frontal Gyrus	65		30	52	32
5	L. Cerebellum Crus I	158		12	-80	-26

Part I: Identifying and modulating neural correlates of PH

	L. Cerebellum Crus II	169		-8	-80	-28
	Inferior Parietal Lobule; Middle Temporal Gyrus; Precuneus; Middle Occipital Gyrus; Superior Temporal Gyrus	14019	7 / 40 / 21 / 39 / 19 / 31	-30	-76	40
	R. Inferior Frontal Gyrus; Medial Frontal Gyrus; Supplemental Motor Area; R. Superior Frontal Gyrus	8929	8 / 10 / 9 / 6 / 46 / 32 / 47 / 45	52	22	28
	L. Middle Frontal Gyrus; L. Inferior Frontal Gyrus; L. Superior Frontal Gyrus	4473	10 / 8 / 9 / 6 / 46	-50	20	28
	R. Postcentral Gyrus	81	3	44	-20	54
6	L. Cerebellum Crus I	134		-8	-78	-26
	R. Posterior Superior Temporal Sulcus (seed)	506		62	-46	-4
	R. Middle Frontal Gyrus (includes IFG seed)	5032	9 / 6 / 46 / 10 / 8 / 45 / 47	52	20	28
	L. Inferior frontal Pars Traingularis	205		-40	42	6
	R. Caudate Body	79		12	2	14
	R. Inferior Parietal Lobule	2374	40	50	-44	52
	L. Inferior frontal Gyrus	208	9	-50	10	36
	L. Inferior Parietal Lobule	902	40	-48	-48	54
	Superior Medial Frontal Gyrus; Supplemental Motor Area	1008	8 / 6	4	34	46
7	L. Cerebellum 6L	307		-30	-68	-26
	R. Posterior Superior Temporal Sulcus (seed)	1579	37 / 21 / 20	62	-48	-12
	R. Inferior Temporal Gyrus	744	37 / 20	-58	-60	-14
	R. Middle Frontal Gyrus (includes IFG seed)	4270	10 / 9 / 46 / 6	52	16	34
	L. Middle Frontal Gyrus	416	10	-42	58	2
	R. Inferior Parietal Lobule	8770	40 / 7 / 2	46	-46	54
	L. Middle Frontal Gyrus	78		-44	34	32
	L. Precentral	157	9	-50	8	36
	Superior Medial frontal Gyrus	53		4	26	50
	L. Superior Frontal Gyrus	260	6	-28	-6	66
	L. Precentral	58		-36	-28	64
8	Superior Temporal Gyrus; Inferior Frontal Gyrus; Insula; Superior Temporal Gyrus; Middle Frontal Gyrus; Middle Temporal Gyrus; Precentral Gyrus; Inferior Parietal Lobule; Middle Temporal Gyrus;	31472	13 / 22 / 40 / 47 / 6 / 21 / 40 / 44 / 9 / 45	-50	-22	6

Part I: Identifying and modulating neural correlates of PH

	Middle Frontal Gyrus; Putamen; Superior Frontal Gyrus				
	Middle Cingulate Cortex; Supplemental Motor Area; Superior Frontal Gyrus	6050	32 / 24 / 6 / 31 / 9	-4	20 32
	R. Precuneus; R. Cuneus	83		16	-66 32
	L. Precentral	55	6	-38	-2 48
9	R. Angular Gyrus	4213	21 / 39 / 40 / 22 / 20	50	-60 28
	R. Cerebellum 6L	279		16	-76 -24
	L. Cerebellum Crus 1	97		-18	-88 -28
	L. Middle Temporal Gyrus (includes pSTS seed)	1614			
	Medial Frontal Gyrus (includes IFG seed)	10912	9 / 10 / 32 / 11 / 6 / 46	0	52 -8
	L. Inferior Frontal Orbital Gyrus	612	47	-46	42 -10
	R. Inferior Frontal Orbital Gyrus	272	47	36	38 -10
	L. Calcarine	52		4	-86 0
	L. Thalamus	7		-4	-14 8
	R. Thalamus	9		8	-14 10
	Precuneus and Posterior Cingulate Cortex	3181	31 / 7 / 23 / 30	4	-52 24
	L. Angular Gyrus	2149	39 / 40	-46	-68 36

Part I: Identifying and modulating neural correlates of PH

Table 2.3 [S3] Occupancy statistics

Z-Statistics for non-parametric Wilcoxon-ranksum test, assessing the effect for the occupancies of the CAPs showed in the rows above. p-value corrected for multiple comparisons in parenthesis.

	Async ≠ Sync	Async ≠ Rest	Sync ≠ Rest
CAP 0	-4.14 (< 0.001)	-1.90 (0.071)	2.70 (0.014)
CAP 1	2.31 (0.051)	2.15 (0.053)	-0.29 (0.770)
CAP 2	0.66 (0.635)	2.54 (0.022)	1.94 (0.074)
CAP 3	0.04 (0.969)	-1.97 (0.069)	-2.25 (0.041)
CAP 4	0.72 (0.670)	0.29 (0.770)	-0.59 (0.614)
CAP 5	-0.87 (0.639)	-5.93 (< 0.001)	-5.70 (< 0.001)
CAP 6	2.44 (0.049)	3.41 (0.002)	1.17 (0.302)
CAP 7	2.09 (0.074)	4.85 (< 0.001)	3.98 (< 0.001)
CAP 8	4.18 (< 0.001)	-0.49 (0.700)	-3.84 (< 0.001)
CAP 9	0.56 (0.642)	4.11 (< 0.001)	4.03 (< 0.001)

Table 2.4 [S4] Average duration statistics

Linear mixed model summary statistics for the effect, on average duration, of each level of the parameter condition, for every CAP. Estimate effects of the synchronous condition, and of rest are shown with respect to the asynchronous condition. The p-value refers to any effect of condition of average duration.

	Condition									
	Async			Sync			Rest			-
	Estimate (seconds)	SE	T	Estimate Effect (w.r.t. to Async)	SE	T	Estimate Effect (w.r.t. to Async)	SE	T	p-value
CAP 0	13.84	0.47	29.67	+2.78	0.65	4.28	-2.06	0.65	-3.16	< 0.001
CAP 1	3.29	0.16	20.00	-0.25	0.23	-1.09	-0.42	0.23	-1.85	0.174
CAP 2	2.91	0.09	32.73	+0.01	0.12	0.05	-0.23	0.12	-1.85	0.097
CAP 3	3.09	0.21	14.90	-0.07	0.24	-0.29	+0.17	0.23	0.72	0.557
CAP 4	2.73	0.08	35.18	-0.09	0.11	0.80	+0.07	0.11	0.65	0.687

Part I: Identifying and modulating neural correlates of PH

CAP 5	3.18	0.30	10.57	-0.42	0.35	-1.21	+0.23	0.32	0.73	0.064
CAP 6	3.56	0.17	20.66	-0.60	0.24	-2.55	-0.57	0.24	-2.33	0.021
CAP 7	3.41	0.16	21.81	-0.30	0.22	-1.39	-0.64	0.22	-2.95	0.015
CAP 8	2.74	0.08	32.51	-0.24	0.12	-1.98	-0.03	0.12	-0.24	0.098
CAP 9	3.01	0.16	18.69	-0.02	0.23	-0.10	+0.04	0.24	0.17	0.97

Table 2.5 [S5] Hierarchical models for transition probabilities

Comparison between different models for the transition probabilities. The model being tested is shown under Model Designation, and the model against it is compared is immediately on the right. For the sake of a more understandable reading, we also highlight here that parameter of interest that is being compared between the models. Degrees of freedom are shown under DoF, and in parenthesis one can see the DoF of the model that the current model is being compared with. Log likelihood is shown also with the same logic. Approximate difference in deviance between the two models is also shown, with a negative value meaning a drop in deviance favouring the main model, over the model it's being tested against. Finally, the p-value shown, is the result of the Wald Chi-Square test.

Model Designation	Tested Against Model	Parameter of Interest	DoF	$\log \mathcal{L}$	$\approx \Delta$ Deviance	p-value
Model 0	-	1 Participant	3	924.47	-	-
Model I	Model 0	CAP _{init}	12 (3)	935.32 (924.47)	-22	0.009
Model II	Model I	CAP _{next}	21 (12)	3274.5 (935.32)	-4679	< 0.001
Model III	Model II	CAP _{init} :CAP _{next}	102 (21)	3788.3 (3274.5)	-1027	< 0.001
F ₉ I	F ₉ 0	Cond	14 (12)	278.3 (272.4)	-11	0.008
B ₆ I	B ₆ 0	Cond	14 (12)	506.15 (495.70)	-21	< 0.001
B ₇ I	B ₇ 0	Cond	14 (12)	621.31 (613.44)	-16	< 0.001
B ₉ I	B ₉ 0	Cond	14 (12)	579.58 (574.50)	-10	0.006
B ₆ III (PH)	B ₆ II (PH)	Cond:PH	17 (15)	512.41 (507.98)	-8	0.036
B ₉ III (PE)	B ₉ II (PE)	Cond:PE	17 (15)	585.78 (581.04)	-9	0.026
B ₆ ^A I (PH)	B ₆ ^A 0 (PH)	PH	13 (12)	116.07 (113.01)	-6	0.039
B ₉ ^A I (PE)	B ₉ ^A 0 (PE)	PE	13 (12)	170.88 (167.57)	-6	0.030

Part I: Identifying and modulating neural correlates of PH

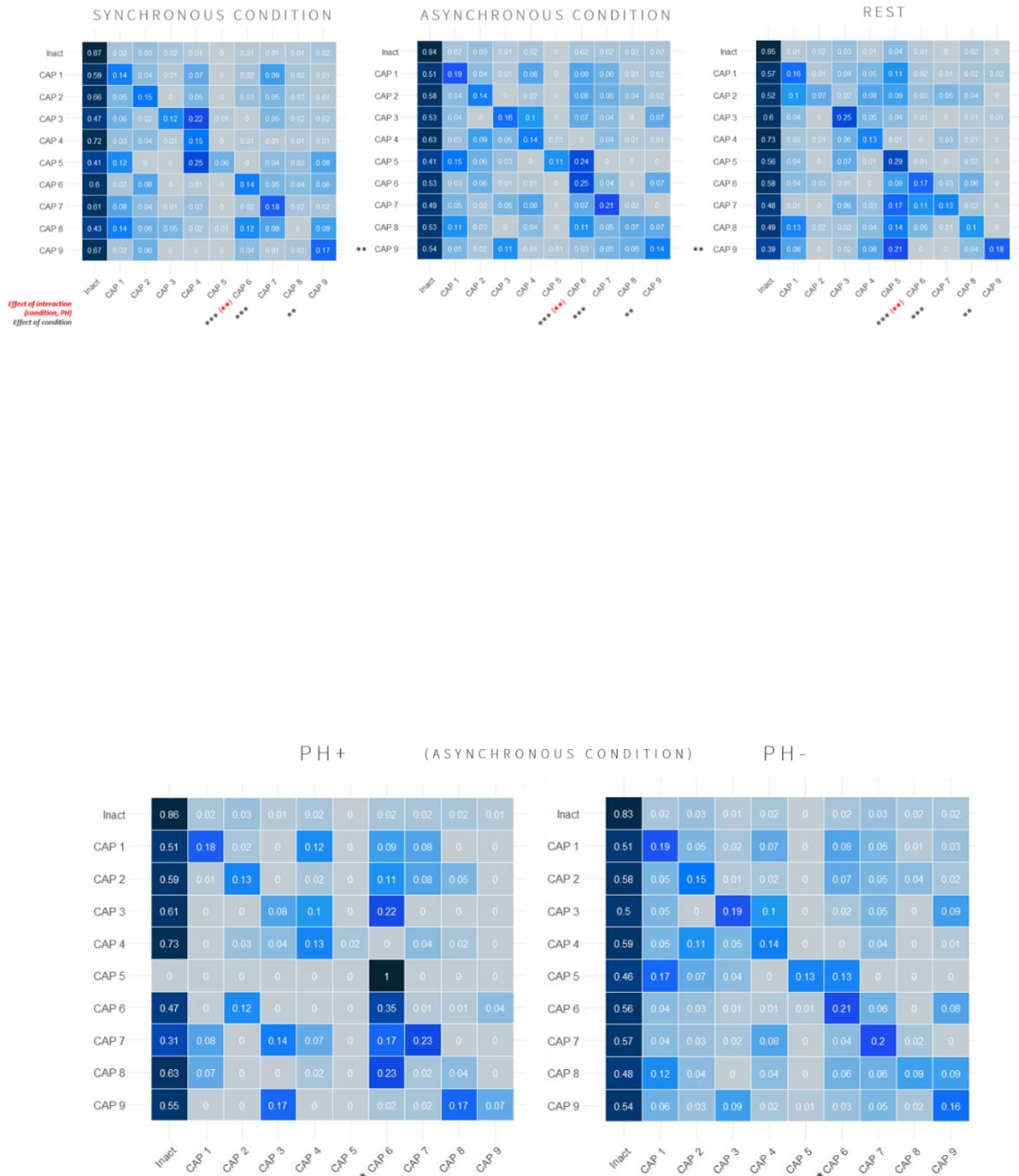


Figure 2.9 [S4] CAP transitions in the asynchronous condition for participants who experienced induced PH and for those who did not

Transition matrixes from all CAPs are shown for the asynchronous condition, for the participants that experienced the induced PH, and for those that did not (measured as subjective difference between asynchronous and synchronous conditions). CAP 6, that showed an interaction between condition and experience PH, was tested for a fixed effect of PH, in every condition. An effect of PH was observed for this brain state in the asynchronous condition (p-val = 0.039).

Part I: Identifying and modulating neural correlates of PH

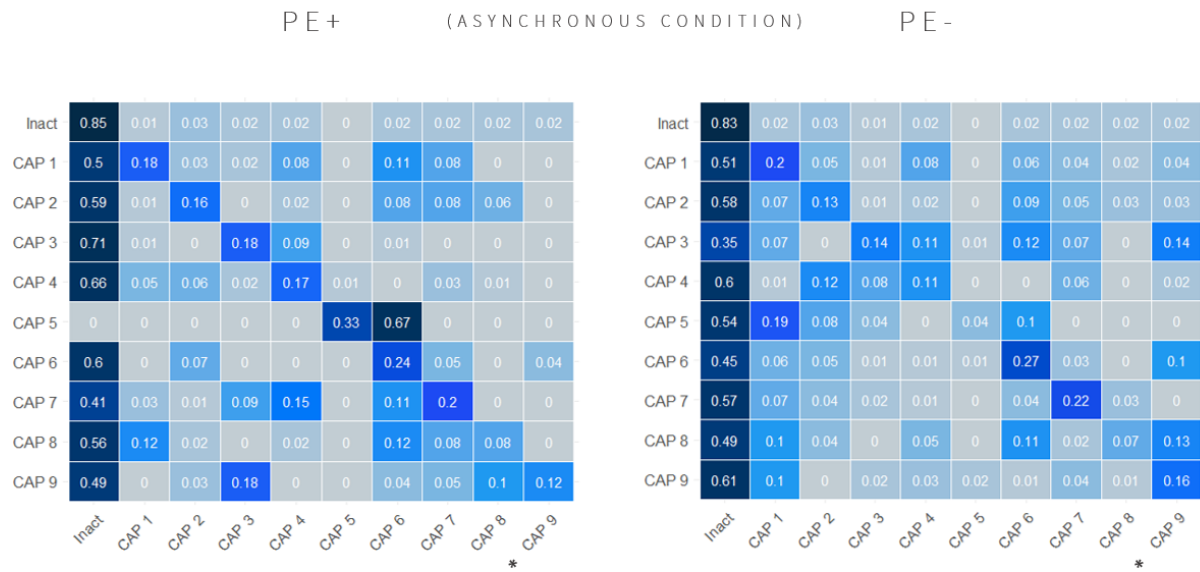


Figure 2.10 [S5] CAP transitions in the asynchronous condition for participants who experienced induced PE and for those who did not

Transition matrixes from all CAPs are shown for the asynchronous condition, for the participants that experienced the induced PE, and for those that did not (measured as subjective difference between asynchronous and synchronous conditions). CAP 9, that showed an interaction between condition and experience PE, was tested for a fixed effect of PE, in every condition. An effect of induced PE was only observed for this CAP in the asynchronous condition (p-val = 0.030).

Part I: Identifying and modulating neural correlates of PH

Control Analyses

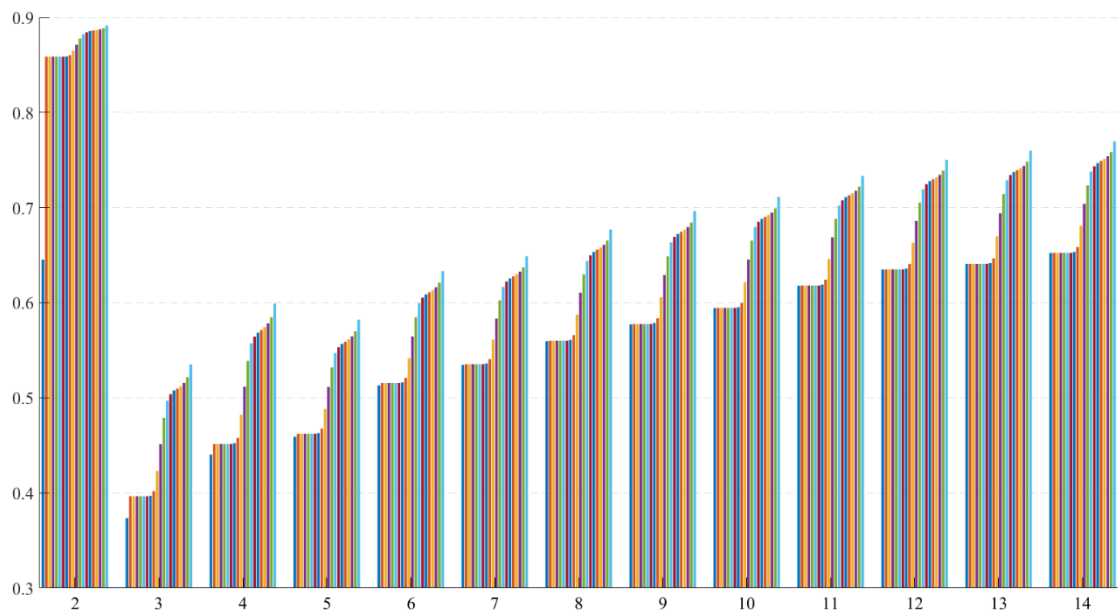


Figure 2.11 [S6] Consensus clustering (control - PCC)

Output from the TbCaps (Bolton, et al. 2020) computed for a different number of centroids (k), for the control analysis with seed on the PCC. The number of centroids was selected at 6.

Table 2.6 [S6] Average duration statistics (control - PCC)

Linear mixed model summary statistics for the effect, on average duration, of each level of the parameter condition, for every CAP, for the control analysis performed with seed on the PCC.

	Condition									
	Async			Sync			Rest			-
	Estimate (seconds)	SE	T	Estimate Effect (w.r.t. to Async)	SE	T	Estimate Effect (w.r.t. to Async)	SE	T	p-value
CAP 0	19.89	0.48	41.85	-0.26	0.67	-0.39	-5.92	0.67	-8.81	<0.001
CAP 1	4.32	0.33	13.29	-0.032	0.43	-0.07	-0.81	0.46	-1.74	0.22
CAP 2	2.73	0.29	9.55	0.10	0.38	0.27	1.13	0.31	3.61	0.002
CAP 3	3.23	0.17	19.38	-0.08	0.23	-0.36	-0.07	0.23	-0.324	0.921
CAP 4	3.26	0.23	14.31	0.42	0.27	-0.97	-0.27	0.28	-0.97	0.109

Part I: Identifying and modulating neural correlates of PH

CAP 5	3.35	0.23	14.45	-0.40	0.27	-1.48	0.07	0.26	0.28	0.224
CAP 6	3.01	0.16	18.34	-0.04	0.19	-0.24	0.22	0.19	1.20	0.320

Control Analyses

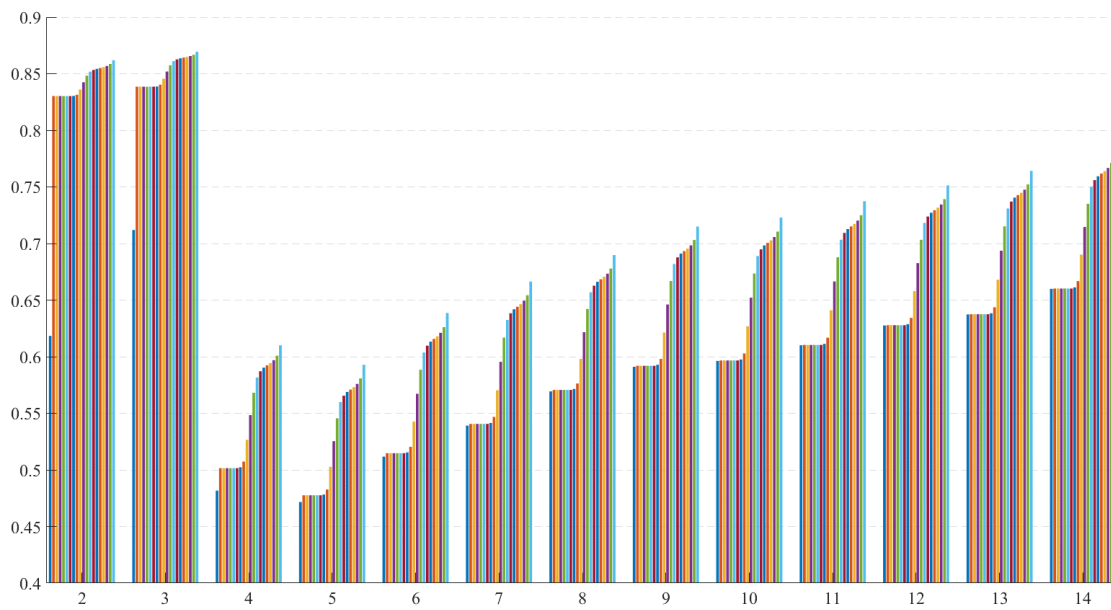


Figure 2.12 [S7] Consensus clustering (control - mSFG and aINS)

Output from the TbCaps (Bolton, et al. 2020) computed for a different number of centroids (k), for the control analysis with seeds on the aINS and mSFG. The number of centroids was selected at 9.

Part I: Identifying and modulating neural correlates of PH

Table 2.7 [S7] Average duration statistics (control - mSFG and aINS)

	Condition									
	Async			Sync			Rest			-
	Estimate (seconds)	SE	T	Estimate Effect (w.r.t. to Async)	SE	T	Estimate Effect (w.r.t. to Async)	SE	T	p-value
CAP 0	14.33	0.58	24.58	2.39	0.69	3.48	-1.92	0.69	-2.78	< 0.001
CAP 1	3.13	0.15	21.08	-0.26	0.21	-1.23	-0.16	0.21	-0.76	0.500
CAP 2	3.81	0.19	20.42	-0.32	0.27	-1.19	-0.98	0.26	-3.71	0.005
CAP 3	4.01	0.30	13.20	-0.93	0.41	-2.25	-1.02	0.41	-2.51	0.067
CAP 4	3.40	0.21	16.17	-0.23	0.30	-0.75	-0.67	0.31	-2.16	0.152
CAP 5	3.11	0.12	25.21	-0.27	0.16	-1.68	-0.12	0.16	-0.76	0.295
CAP 6	4.20	0.39	10.84	-0.71	0.49	-1.44	-1.19	0.47	-2.53	0.086
CAP 7	3.41	0.16	21.22	-0.36	0.22	-1.64	-0.31	0.23	-1.33	0.294
CAP 8	2.90	0.23	12.73	-0.86	0.28	-0.31	0.69	0.22	3.18	0.005
CAP 9	3.06	0.36	8.52	0.46	0.46	1.01	0.28	0.41	0.69	0.575

Part I: Identifying and modulating neural correlates of PH

Table 2.8 [S8] Comparison between different models for transition probabilities (control - mSFG and aINS)

Model Designation	Tested Against Model	Parameter of Interest	DoF	$\log \mathcal{L}$	$\approx \Delta$ Deviance	p-value
Model 0	-	1 Participant	3	894.97	-	-
Model I	Model 0	CAP _{init}	12 (3)	906.70 (894.97)	-23	0.005
Model II	Model I	CAP _{next}	21 (12)	3137.4 (906.70)	-4461.5	< 0.001
Model III	Model II	CAP _{init} :CAP _{next}	102 (21)	3728.9 (3137.4)	-1182.9	< 0.001
F ₂ I	F ₂ 0	Cond	14 (12)	456.73 (456.65)	-0.15	0.93
B ₂ I	B ₂ 0	Cond	14 (12)	358.37 (355.78)	-5.17	0.07
F ₂ III (PH)	F ₂ II (PH)	Cond:PH	17 (15)	457.07 (456.77)	-0.61	0.74
B ₂ III (PH)	B ₂ II (PH)	Cond:PH	17 (15)	360.43 (359.72)	-1.40	0.50
F ₂ III (PE)	F ₂ II (PE)	Cond:PE	17 (15)	456.98 (456.74)	-0.49	0.78
B ₂ III (PE)	B ₂ II (PE)	Cond:PE	17 (15)	360.05 (358.39)	-3.31	0.19

Supplementary References

Bolton, T.A.W., Tuleasca, C., Wotruba, D., Rey, G., Dhanis, H., Gauthier, B., Delavari, F., Morgenroth, E., Gaviria, J., Blondiaux, E., Smigielski, L., Van De Ville, D., 2020b. TbCAPs: A toolbox for co-activation pattern analysis. *Neuroimage* 211, 116621. <https://doi.org/10.1016/j.neuroimage.2020.116621>

Monti, S., Tamayo, P., Mesirov, J., Golub, T., 2003. Consensus clustering: A resampling-based method for class discovery and visualization of gene expression microarray data. *Mach. Learn.* 52, 91–118. <https://doi.org/10.1023/A:1023949509487>

2.2 Study 2: Sustained bidirectional self-regulation of hallucination networks through real-time fMRI neurofeedback

Herberto Dhanis ^{1,2,4}, Nicolas Gninenko ^{1,4}, Elenor Morgenroth ^{1,4}, Jevita Potheegadoo ^{1,2},
Giulio Rognini ^{1,2}, Nathan Faivre ^{1,2,5}, Olaf Blanke ^{1,2,6,a,*}, and Dimitri Van De Ville ^{1,4,a,*}

¹ Neuro-X Institute, Ecole Polytechnique Fédérale de Lausanne, EPFL, Geneva, Switzerland

² Brain Mind Institute, Faculty of Life Sciences, Ecole Polytechnique Fédérale de Lausanne, Lausanne, Switzerland

⁴ Department of Radiology and Medical Informatics, University of Geneva, Geneva Switzerland

⁵ Univ. Grenoble Alpes, Univ. Savoie Mont Blanc, CNRS, LPNC, 38000 Grenoble, France

⁶ Department of Clinical Neurosciences, University Hospital of Geneva, Geneva, Switzerland

^a *Both authors contributed equally*

^{*} *Corresponding authors*

Corresponding authors

Prof. Olaf Blanke
Bertarelli Chair in Cognitive Neuroprosthetics
Neuro-X Institute & Brain Mind Institute
School of Life Sciences
Ecole Polytechnique Fédérale de Lausanne (EPFL)
Campus Biotech, H4.3
Ch. des Mines 9
CH-1202 Geneva
E-mail: olaf.blanke@epfl.ch
Tel: +41 (0)21 693 69 21

Prof. Dimitri Van De Ville
Neuro-X Institute
School of Engineering
Ecole Polytechnique Fédérale de Lausanne (EPFL)
Campus Biotech, H4.3
Ch. des Mines 9
CH-1202 Genève
E-mail: dimitri.vandeville@epfl.ch
Tel: +41 (0)21 693 96 69

2.2.1 Abstract

Real-time fMRI neurofeedback (fMRI-NF) can modulate resting-state markers of neuropsychiatric symptoms, but has not yet been shown to modulate neural networks during ongoing hallucinations, nor whether such modulation has lasting effects. Here, we combined fMRI-NF and co-activation pattern analysis with MRI-compatible robotics capable of inducing a clinically-relevant hallucination, to train healthy individuals to modulate a brain network associated with the induced hallucination. Using bidirectional fMRI-NF participants learned, over three days, to successfully up- and down-regulate the dynamics of the hallucination network, which lead to an increase in sensitivity to the robotically-induced hallucination post-training. Furthermore, newly sensitive participants that succeeded in fMRI-NF presented sustained neural changes post-training, specifically, increased hallucination network occurrences during induction and decreased during a matched control condition. Neuroscience-robotics for hallucinations paired with fMRI-NF modulate the dynamics of hallucination networks, alter hallucinatory experience and underlying brain dynamics, which could have direct translational relevance for novel antipsychotic therapies in disease.

Keywords: Real-time fMRI, Neurofeedback (NF), Co-Activation Pattern (CAP), Dynamic Functional Connectivity, Temporal Processes

2.2.2 Introduction

Hallucinations are complex and heterogeneous phenomena during which an individual has an aberrant perceptual experience in the absence of any corresponding external stimulus ¹. While hallucinations can be experienced by healthy individuals ², with increased prevalence in the elderly ^{3,4} and during bereavement ^{3,5}, they are of major clinical relevance in psychiatric conditions such as schizophrenia ⁶ or neurodegenerative diseases such as Lewy-Body Dementia ⁷ and Parkinson's Disease ^{8,9}. Despite this, hallucinations remain difficult to study in laboratory conditions, given their spontaneous appearance, variable content, and lack of specific induction protocols. Methods such as Ganzfeld ¹⁰ and flicker-induced hallucinations ¹¹ have been developed for this purpose, but do not induce clinically relevant hallucinations (i.e., similar to symptomatic hallucinations experienced by patients) and lack good experimental control over their induction (real-time induction; induction across controlled conditions). Psychoactive drugs can induce vivid hallucinations, but these are often unspecific, have highly variable content, and are often accompanied by significant alterations of consciousness ^{12,13}.

Recently, the integration of robotics technology with insights from cognitive neuroscience and clinical research, allowed the repeated induction of a specific and clinically-relevant hallucination: presence hallucination (PH), which is the convincing sensation that someone is close by when no one is there ¹⁴. PH is frequent in schizophrenia ¹⁵, Lewy-Body Dementia ^{16,17}, Parkinson's Disease ¹⁸, and has also been reported in stroke ¹⁹, epilepsy ²⁰, as well as healthy individuals ²¹. PH is one of the most common hallucinations in PD (70% of patients) ¹⁸, occurring early in the disease ²², and is associated with faster disease progression leading to structured hallucinations, persistent psychosis, and cognitive decline ^{23,24}, which in turn are associated with higher risk of mortality ^{25,26}. The PH-inducing robotic setup has been shown to induce PH in healthy individuals ^{27–31}, first-episode psychosis patients ³² and in patients with Parkinson's Disease ³³, by exposing them to different sensorimotor conflicts. Specifically, participants manipulate a front-robot, whilst a back-robot provides tactile feedback to their back, with a temporal delay, by reproducing the participant's front movements. Such asynchronous stimulation triggers robot-induced PH (riPH; PH-inducing condition), whereas a control condition without the temporal delay but otherwise identical stimulation, does not (control condition). Recently, this robotic setup was extended to the MRI-setting ³¹ in healthy individuals ^{33,34}, and revealed that both the posterior superior temporal sulcus (pSTS) and the inferior frontal gyrus (IFG) are more active in the PH-inducing condition rather than in the control condition ³³. Further investigation of brain networks dynamics during periods of high activity of the pSTS and IFG, identified that the brain networks of participants who were sensitive to riPH, increased their transitions to a specific PH-state ³⁵.

Given the clinical relevance of PH ^{18,24}, the ability to induce PH in the MR-scanner and the identification of abnormal temporal dynamics related to the PH-state in healthy individuals, here, we investigated (1) whether healthy participants could achieve volitional control of the PH-state, (2) whether successful regulation would increase sensitivity to riPH, and (3) whether such regulation would trigger changes in its underlying brain mechanisms. Together, such findings would allow a more causal inference of the role of this brain state in PH, and could potentially pave the way for therapeutic antipsychotic fMRI studies, for example in patients with Parkinson's Disease. To that aim, we combined the MR-compatible PH-inducing robot ³¹ with real-time fMRI neurofeedback (NF) ³⁶.

Part I: Identifying and modulating neural correlates of PH

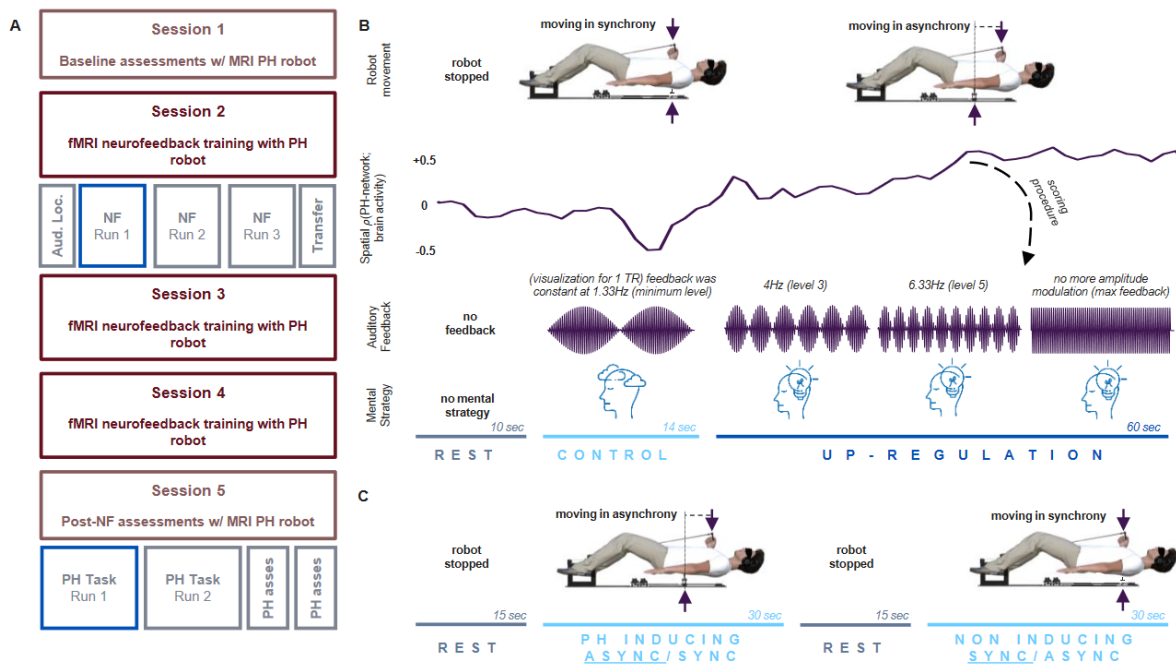


Figure 2.13 Experimental Paradigm

A) Diagram of the fMRI experiment. On average participants completed the experiment in around two weeks, with a minimum break of one day between each session. **(B)** One block of neurofeedback training is shown. In total six were done per run. Participants start at rest. After one auditory cue the control condition starts and participants move the robot in the synchronous mode, while listening to the minimal level of auditory feedback (it did not vary with brain activity at this point). Participants were told they shouldn't engage in the up-regulation strategy here and should rather "clear their minds". After two auditory cues the up-regulation condition started, and participants continued moving the robot which changed to the asynchronous mode. The auditory feedback (resembled a looming and receding wave) became now variable with the spatial correlation between ongoing brain activity and the PH network. Participants knew they should deploy a mental strategy to boost the feedback to the maximum level. This condition ended with three auditory cues. Participants were asked about preferred strategies at the end of each session (and reassured some strategies might not be verbalizable). At the end of session 4 a more general briefing occurred. **(C)** Two blocks of the PH induction sessions are shown in sequence of a total of 16 blocks per run. Blocks were composed of rest followed by an auditory cue which indicated the start of the condition where participants moved the robot in asynchronous or synchronous mode. Two auditory cues indicated the end of the PH task. No more than two blocks in sequence could be of the same condition. At the end of the second run of PH induction session, participants performed two individual blocks of robot manipulation (one asynchronous, one synchronous), each followed by a questionnaire assessing PH induction sensitivity amongst other sensations and control questions.

Real-time fMRI NF is a technique to "close the loop" between brain activity and behaviour, by providing participants with real-time feedback about specific features of their brain activity, with the aim of enabling the search for a mental strategy capable of successfully regulating said brain activity and modulate behaviour^{36–38}. Using NF to target large-scale functional brain networks comparable to the PH-state, has been demonstrated for attention³⁹ and stress⁴⁰. Targeting and controlling functional connectivity through NF is promising in translating such control of brain activity to modulation of neuropsychiatric symptoms^{37,41,42,43}. For the present experiment we developed a paradigm where participants performed five fMRI sessions (each on a different day), spanning over the course of two weeks (Fig. 2.13A): one PH-induction session for pre-training rIPH assessments, three sessions of NF-training targeting the PH-state, and one more PH-induction session for post-training rIPH assessments. During NF-training sessions, participants manipulated the PH-inducing robot, while attending a continuous auditory feedback score based on the correlation between their ongoing brain activity and

the PH-state (NF-signal). The auditory feedback only varied with the NF-signal during the PH-NF condition, and was otherwise at minimum level during control-NF (Fig. 2.13). Hence, participants were instructed to develop a mental strategy to achieve the highest level of feedback during the PH-NF condition (for details see methods; participants were blinded to the purpose of NF). With this, we trained 20 participants to up-regulate the PH-state in a total of 176 NF-training runs, assessed their regulation performance (Fig. 2.14), and investigated ongoing brain activity. Our results show that NF runs, during which participants achieved significant regulation of the PH-state, were reflected by two different neural mechanisms: up-regulation through re-instantiation of the PH-state during PH-NF and down-regulation through avoidance of the PH-state during control-NF. Three days of NF-training, led to an increase in sensitivity to riPH after training. Moreover, participants who became sensitive to riPH, presented higher occurrences of the PH-state during PH-induction, and decreased occurrences during a control condition.

2.2.3 Results

2.2.3.1 Neurofeedback Regulation Performance

We first investigated whether there was any evidence for up- and/or down-regulation of the PH-state. This was done by examining the auditory feedback scores, given that these scores represented activity of the PH-state during PH-NF normalised by the median activity during the previous control-NF (see Methods), and thus indicated instances of PH-state modulation between the two conditions (Fig. 2.14A). Hence, we used a non-parametric surrogate-data procedure, to generate a null distribution of the average PH-NF feedback scores for each NF run (Fig. 2.14B). We considered a NF run to exhibit high-regulation performance (HRP) when the average feedback score was above a specific percentile threshold of the NF run's surrogate distribution feedback score. This selection threshold was set at 84.1% (i.e., one standard deviation above the mean for a normal distribution) and revealed a total of 52 out of 176 NF runs that achieved HRP (~30%; median per participant was 2; MAD:1.7; range 0-6; $P < 0.00001$) (Fig. 2.14C). Repeating this procedure for every selection threshold between 0.5 and 0.975 (in steps of 0.025), we confirmed that this modulation of the PH-state between PH-NF and control-NF was significant regardless of the chosen threshold (Fig. 2.14D). To further evaluate regulation performance, we sought to assess if HRP was maintained, in runs during which participants received no feedback about regulation success, but that were identical to NF runs in all other aspects (transfer runs, performed at the end of each NF session) (Fig. 2.13A). Using the procedure described above, we identified 18 HRP transfer runs out of 56 (~32%; median per participant: 1; MAD: 0.63; range: 0-2; $P < 0.00001$), and importantly, showed that the number of HRP transfer runs achieved was significantly correlated to the number of HRP runs during the NF sessions ($\rho = 0.49$; $P = 0.02750$; Fig. 2.15). These results show that participants successfully achieved control over the PH-state (HRP) during NF-training, and that they were able to maintain such control even in the absence of feedback (i.e. during transfer runs), which in turn suggests that participants effectively learned and integrated the strategies to regulate the PH-state, while using an MRI-compatible robotic device to induce PH ³³.

Part I: Identifying and modulating neural correlates of PH

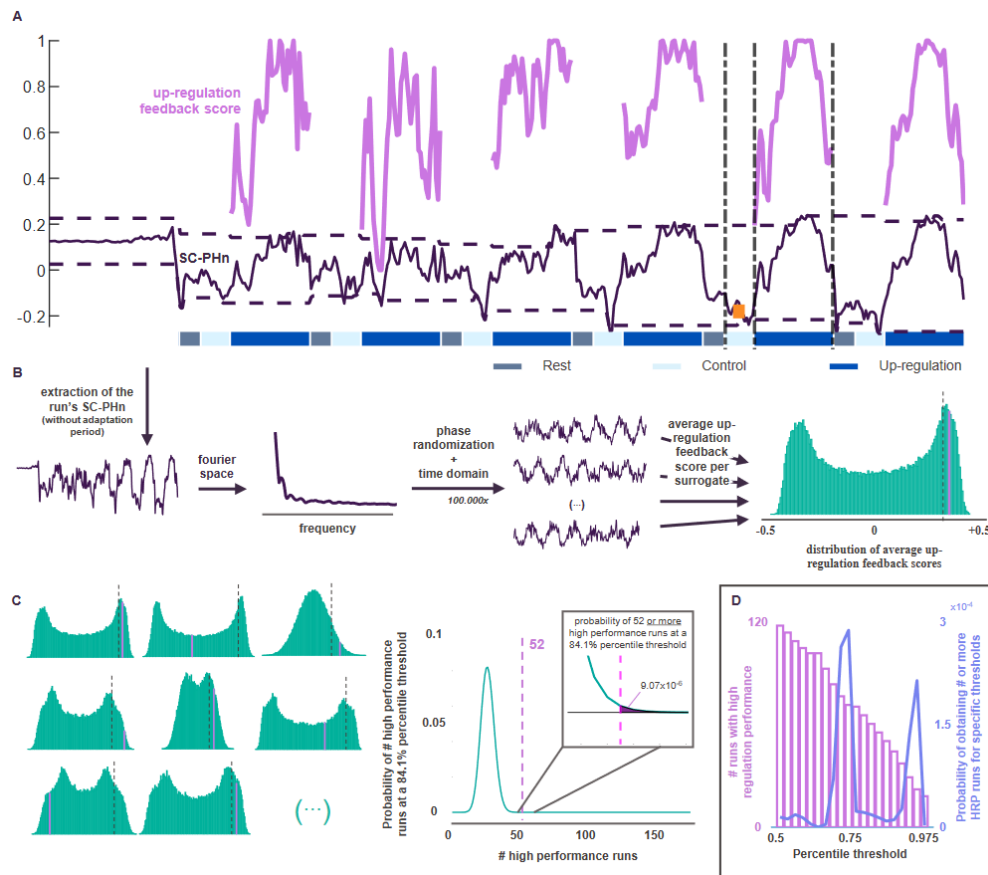


Figure 2.14 Assessment of Neurofeedback Regulation Performance

(A) An exemplary run illustrates how the feedback score presented in up-regulation is calculated. Dynamic range is shown in horizontal dashed lines. Vertical dashed lines highlight the control and up-regulation conditions of one block, with the orange square showing the median SC-PHn of that control period. Computing the feedback score of a time point in up-regulation is performed by normalising the SC-PHn by that timepoint's dynamic range, and subtracting the median SC-PHn of the previous control period (also normalised by the dynamic range). (B) Non-parametric bootstrapping procedure to identify runs with high regulation performance. The SC-PHn of a run is extracted and transformed into Fourier space. Surrogates with identical temporal properties are then generated, by randomising the phase and performing an inverse transformation back into time domain, 100,000 times. By virtually computing the feedback score of each surrogate, a null distribution of average up-regulation feedback scores of the surrogates is obtained, and the observed feedback score can then be compared against this null distribution based on a 84.1% percentile threshold. (C) Examples of null distributions of different runs are shown on the left. On the right, it is shown how the probability of obtaining the observed number of high regulation performance runs (52) or more, was obtained, by evaluating the area under the binomial distribution curve with parameters $X \sim \text{Bin}(176, 1 - 0.841)$. (D) The number of high regulation performance runs is shown on the left, in function of the choice of threshold. On the right the probability of obtaining such number (or more) of high regulation performance (HRP) runs, considering the different percentile thresholds, is shown. Independently of the chosen threshold, regulation was successful in such a manner that the probability of obtaining the observed numbers of high regulation performance runs is always extremely low, even when considering null hypotheses with extremely stringent percentile thresholds.

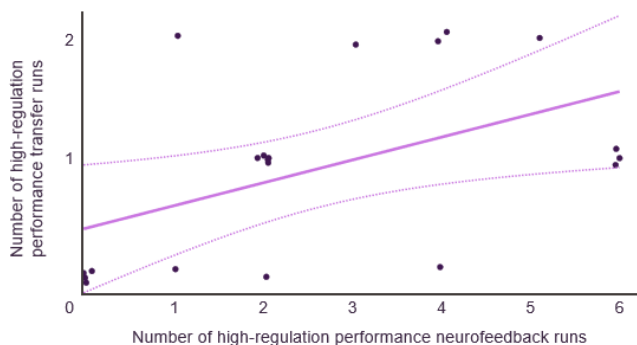


Figure 2.15 High-regulation performance in neurofeedback and transfer runs

The number of HRP runs achieved by participants during neurofeedback training significantly correlated with the number of HRP runs achieved during transfer runs. Given that during transfer runs participants have no feedback informing them of the success of their modulation, this suggests that participants learned the regulation mechanisms they were using, the more successful was their regulation during neurofeedback training.

Data was jittered for visualisation purposes only.

2.2.3.2 Brain States Underlying NF and PH-Induction Sessions

In parallel, we investigated brain states active during NF-training and PH-induction sessions, independently from the HRP assessment, as to allow analyses coupling these data together. Recovering active brain states was performed using co-activation pattern analysis (CAPs)⁴⁴, a clustering method that resolves dynamics of FC in terms of large-scale activation patterns that occur when relevant seed regions exhibit high activity (see Methods). We focused on periods of high activity of two regions related to PH-induction (pSTS and IFG)³³ that have previously been associated with the PH-state³⁵. Periods when these regions were not active are referred to as the inactive state (state⁰). Using this method on runs of NF-training and pre/post training PH-induction sessions, we recovered eight active brain states. These were named based on their topography and included, the PH-state (Figs. 2.16A-C), the sensorimotor+ state (Fig. 2.16D), its opposite pattern (i.e., sensorimotor- state) (Fig. 2.16E), and a DMN- state characterized by deactivated regions of the default mode network⁴⁵ (Fig. 2.16F). We note that the sensorimotor- and the DMN- states had large overlapping activations with the PH-state. All other states are shown in the Extended Data Fig. 1 (cluster peaks in Supplementary Table 1). These findings corroborate previous results that recovered the PH-state during PH-induction, but extend them to the PH-state during NF.

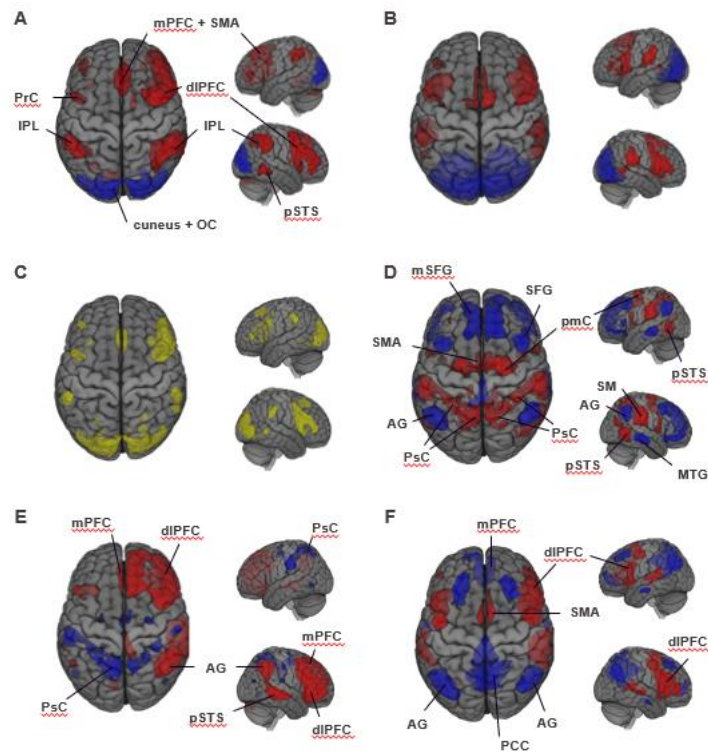


Figure 2.16 Anatomy of the PH-network and of other identified brain states

All brain states are shown after spatial standardization. (A) PH-network identified in previous work³⁴ and used here to compute online SC-PHn. Major clusters are seen on the the right posterior superior temporal sulcus (pSTS), the bilateral inferior parietal lobule (IPL) with focus on the supramarginal gyri (SMG), the right dorso-lateral prefrontal cortex (dlPFC), the middle prefrontal cortex (including part of the supplemental motor area: SMA), the left dlPFC, the left precentral gyrus (PrC), the body of the caudate on the right, and in the cerebellum, the left Crus I and II. Deactivations are observed over the cuneus and occipital gyrus. (B) PH-state as identified by CAP analysis in the neurofeedback data. (C) Overlap between the PH-network³⁴ and the PH-state, showing that the PH-state recovers every major cluster of the PH-network³⁴. (D) The sensorimotor+ state represents mostly sensory processing. Major activation clusters can be seen bilaterally over the sensory cortex (post central gyri, PsC), SMA, pre-motor cortex (pmC) pSTS, and supra-marginal gyri (SM). Major cluster deactivations are also seen bilaterally for the middle temporal gyri (MTG), medial and superior frontal gyri (mSFG, SFG), angular gyri, and posterior cingulate cortex (PCC). (E) The sensorimotor- state is characterized by the pairing of bilateral deactivations of the sensory cortex, and by activations over four main cluster of the PH-state all on the right: the pSTS, the dlPFC, the mPFC, and the AG. (F) The DMN- state is characterized by deactivations of the PCC, bilateral AG, bilateral superior frontal gyri, and mPFC. It is also accompanied by activations over the PH-state clusters, including: SMA, and dlPFC (more prominently on the right).

2.2.3.3 Re-Instantiation of the PH-network During Up-regulation

Next, in order to characterize the brain mechanisms underlying HRP, we computed state's transition and occurrence probabilities, which had been previously associated with PH-induction³⁵. These were analysed with Linear Mixed Models (LMM) followed by post-hoc analyses to decompose significant interactions (multiple comparison corrected, see methods). Our results identified two distinct brain regulation mechanisms underlying HRP. An up-regulation mechanism based on changes in transition probabilities and characterized by re-instantiation of the PH-state during PH-NF, and a down-regulation mechanism based on changes in occurrences and characterized by avoidance of the PH-state during control-NF.

Part I: Identifying and modulating neural correlates of PH

We hypothesized that successful participants would learn to boost the auditory feedback score by changing transition probabilities across states, in a manner that favoured the PH-state. To test this, we first demonstrated that succession probability changed across all brain states (transitions coded with two predictors, initial and final state; see methods), conditions (PH-NF, control-NF, and rest), and depending on whether HRP was achieved (three-way interaction: $P = 0.00058$, Supplementary Table 2). Then, through post-hoc analyses (detailed on Supplementary Tables 3 and 4), we showed that achieving HRP modulated the transitions departing from specific states (to all states), and that furthermore this HRP effect was also dependent on the final state of the transition (interaction between HRP and final state). The initial states which exhibited these changes were: state0 ($P < 0.00001$, $P = 0.01371$, respectively for HRP and interaction), the PH-state ($P < 0.00001$, $P < 0.00699$), the sensorimotor- state ($P < 0.00001$, $P = 0.00036$), the DMN- state ($P < 0.00001$, $P = 0.00047$), and the DMN+ state ($P < 0.00001$, $P = 0.01910$). These significant changes in transition probabilities effectively presented as increases in transition probabilities towards the PH, sensorimotor-, and DMN- states (Fig. 2.17, Supplementary Fig. 3). The same analysis performed for transfer runs (i.e., no auditory feedback) did not recover the exact same pattern, but did detect an increase in transitions from the PH-state to the DMN- (see Supplementary: Tables 5, 6, Fig. 4). While this shows that participants did not fully retain the up-regulation strategy in the absence of feedback, partial retention for the PH-state was found.

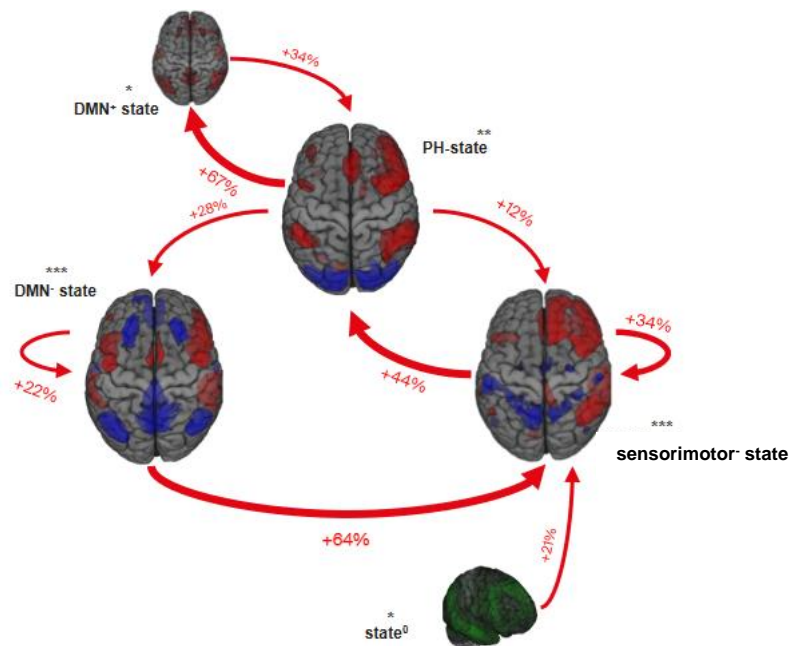


Figure 2.17 Re-instantiation mechanism during the up-regulation NF-condition of HRP runs

Linear mixed model analysis of states' transition probabilities identified a triple interaction between condition, high-regulation performance, and initial state of the transition (p -value < 0.001). Subsequent analysis revealed that in the up-regulation condition of runs where high-regulation performance was achieved, three states (PH, DMN-, sensorimotor-) significantly changed their transitions to other states, as compared to the up-regulation condition in runs where high-regulation performance was not achieved. These changes in transition probabilities formed a reinforcement loop to the PH-state through states that were similar to it. They were predominantly marked by increases in the transitions from the PH-state, to two states which include activations over major PH-state clusters (DMN- by 28%, and sensorimotor- by 12%), by increases from the DMN- to itself (22%) and to the sensorimotor- (64%), and of the sensorimotor- to itself (34%) and back to the PH-state (44%).

Part I: Identifying and modulating neural correlates of PH

Overall, these results demonstrate that successful participants managed to regulate their succession of brain states differently than unsuccessful participants, in a manner that benefitted the PH-state, and two other similar states (sensorimotor- and DMN-).

2.2.3.4 Avoidance of the PH-network During Control

Concerning occurrences of brain states, we hypothesized that the PH-state would increase occurrences during PH-NF of HRP runs. To test this we first demonstrated that occurrences changed depending on condition, brain state and HRP (triple interaction: $P = 0.00004$, Supplementary Table 7). Post-hoc analyses showed this effect was driven by a significant decrease in occurrences of the PH ($P = 0.02603$; Fig. 2.18A) and posterior states ($P = 0.01837$, Extended Data Fig. 2), and by a significant increase in occurrences of the sensorimotor+ state ($P = 0.04763$; Fig. 2.18B), when comparing the control-NF condition of HRP and non-HRP runs (Supplementary Tables 8 and 9). Thus, in the control-NF condition participants avoided the recruitment of the PH-state, and rather recruited the sensorimotor+ state, which included more prominent activations in sensory regions as well as deactivations in regions overlapping with the PH-state. While not initially hypothesised, avoidance of the PH-state and its partial deactivation through the sensorimotor+ state during control-NF, lowered the baseline for the next PH-NF, thereby increasing the feedback participants received during the next PH-NF condition. For results on the transfer runs see Supplemental Information (Supplemental Tables 10 and 11; Supplementary Fig. 4). Collectively, we observe that participants leveraged brain dynamics in two distinct ways. First, during PH-NF participants increased auditory feedback by modulating transitions across brain states, effectively forming a “loop” that up-regulated the PH-state and similar states. Second, participants increased auditory feedback by down-regulating the PH-state during the control-NF condition. With this, participants achieved bidirectional control of the feedback.

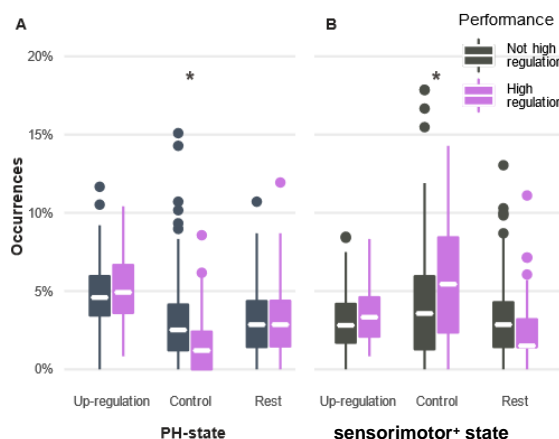


Figure 2.18 Avoidance mechanism during the control condition of HRP runs

(A) Linear mixed model analysis of state's occurrences identified a triple interaction between condition, HRP, and state (p -value < 0.001). Subsequent analysis revealed that the occurrences of the PH-state were significantly lower in the control condition of HRP runs, as compared to the same condition in runs that did not achieve HRP (p -value < 0.05). (B) Subsequent analysis of the triple interaction reported above also revealed that the occurrences of the sensorimotor+ state were significantly higher in the control condition of HRP runs, as compared to the same condition in runs that did not achieve a high-regulation performance (p -value < 0.05).

2.2.3.5 Effects of NF Training on Behavioral Sensitivity to PH Induction

As participants were trained to control the activity of the PH-state, we hypothesized that such NF-training would increase the intensity of riPH post-training. For this, we compared riPH at baseline (day 1) with riPH after NF-training (day 5). On day 1, only 10% of participants (2/20) were sensitive to riPH, as measured by a positive difference in PH ratings between the PH-inducing and control conditions. At the group-level, there was no significant riPH difference between both conditions (as assessed with cumulative linear models, see Methods). This changed after NF-training (day 5), as 40% of participants became sensitive to the robotic stimulation (8/20, including the two already sensitive at day 1), and a significant riPH difference at the group level ($P = 0.01126$; Fig. 2.19) was observed. Importantly, no changes (between day 1 and 5) were observed for other sensations that accompany riPH (passivity experiences: $P = 0.9095$, Extended Data Fig. 3; loss of agency: $P = 0.20910$, Extended Data Fig. 4) nor for control questions (Supplemental Fig. 5). In line with our hypothesis, we show that NF-training targeting the PH-state, in fully blinded participants, increased the participants' riPH sensitivity, but did not alter the sensitivity of accompanying and control sensations. This shows that NF-training combined with sessions during which a specific hallucination was induced (riPH) changed participants' proneness to experiencing that hallucination.

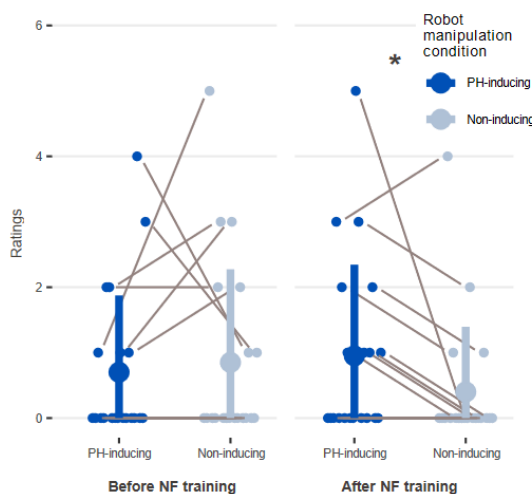


Figure 2.19 PH ratings before and after NF training

Raw ratings for the PH induction question are shown for both asynchronous and synchronous robot manipulation condition, before neurofeedback training (left) and after neurofeedback training (right). Cumulative linear mixed model analysis showed no statistical difference between the conditions before training. After completing neurofeedback training, there was a significant difference ($p\text{-value} < 0.05$) between the two conditions, and more participants were sensitive to the induction of PH. *Data was jittered for visualization.*

2.2.3.6 HRP Changes Brain Activity of Participants Sensitivity to PH in a Directional Consistent Manner with Regulation Strategies

The present NF-training targeting the PH-state led to the identification of a bidirectional neural regulation mechanism (up-regulation through re-instantiation of the PH-state during PH-NF; down-regulation through avoidance of the PH-state during control-NF), and a behavioural increase in riPH sensitivity post-training. Next, we investigated whether the present NF-training also led to differences in the brain dynamics of PH-induction by comparing the session before NF-training (Day 1) with the one after NF-training (Day 5). We used a similar temporal modelling analysis as done for NF-training, but now modelling the differential states' occurrence and transition probabilities between Day 5 and Day 1, and taking into account participants' sensitivity to riPH after NF-training, as well as the number of HRP runs obtained

Part I: Identifying and modulating neural correlates of PH

during NF-training. This, revealed that occurrences changed after NF-training depending on brain state, on how many HRP runs the participants achieved during NF-training, and on their sensitivity to riPH (triple interaction $P = 0.00125$, Supplementary Table 12). Thus, participants, who became sensitive to riPH by NF-training, changed the occurrences of the PH-state differently for the PH-inducing versus control conditions, depending on the number of HRP runs achieved ($P = 0.03555$; Fig. 8A; Supplementary Tables 13 and 14). This interaction indicated that the more successful a participant was during NF-training, the more PH-state occurrences were observed in the PH-inducing condition after NF-training, and the less PH-state occurrences were observed in the control condition. Critically, this effect was absent for participants who did not become sensitive to riPH after NF-training (Fig. 2.20B). Further control analyses, repeating the same procedure with other sensations tested during riPH (passivity experiences and loss of agency), did not show any specific changes related to NF-training (Supplemental Tables 16 to 20). In sum, participants who became sensitive to riPH retained the strategies and neural mechanisms developed during NF training, the more successful they were during NF training. This happened in a bidirectional manner with the re-instantiation of the PH-state, observed for the PH-inducing condition as increased occurrences, and the avoidance of the PH-state during control-NF, observed here for the control condition as decreased occurrences. The same type of analysis for differential transition probabilities analysis did not show a significant differences associated with NF-training (see Supplemental Tables S21 to 26).

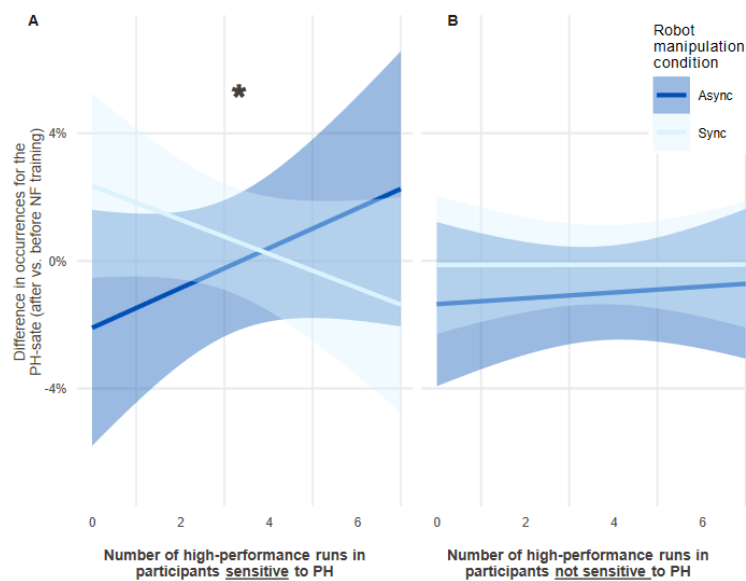


Figure 2.20 Difference in PH state occurrences when comparing after vs. before training

The difference in occurrences of the PH-state for after vs. before neurofeedback training, is shown for each condition respectively, in function of the number of achieved high regulation performance runs. A significant interaction between PH sensitivity, state and number of high-regulation performance runs was observed (p -value < 0.001) (A) This difference is shown for participants that were sensitive to PH induction at the end of training. A significant interaction between number of high-regulation performance runs and condition was observed (p -value < 0.05). (B) Same difference shown for participants that were not sensitive to PH induction on the last day. No interaction is observed.

2.2.4 Discussion

Combining MR-robotics able to induce a clinically-relevant hallucination (PH) ^{31,33}, and real-time fMRI NF, we demonstrated that volitional control over a prototypical brain state of PH is achievable. Specifically, we showed that participants were able to both re-instantiate the PH-state during PH-NF, and avoid it during control-NF. This led to an increased sensitivity to riPH post-training. Furthermore, for participants who became sensitive to riPH following NF-training, the more successful they were in regulating their brain activity, the more they retained the regulation mechanisms after training.

Real-time fMRI NF can train participants to achieve volitional control over hypothesised neural markers of clinical symptoms such as pain ⁴⁶ or PD tremor ⁴⁷. In particular, different studies have shown considerable promise of functional connectivity based NF in modulating neural markers of neuropsychiatric symptoms ^{37,41,42}. We designed a NF paradigm to modulate the brain activity underlying a clinically-relevant hallucination, the PH-state, while PH was robotically induced by our MRI compatible system. To control for confounding factors, such as motivation, difficulty, and differences in manipulation of the robot across participants, as well as placebo, and spatially unspecific effects, PH-NF feedback was normalized using a baseline established in every preceding control-NF condition ⁴⁸. Ultimately this allowed participants to develop bi-directional regulation strategies, by up-regulating the PH-state during the PH-NF condition and by down-regulating it during the control-NF condition.

During the PH-NF condition of HRP runs, we observed a rearrangement of transitions across brain states, which favored both the PH-state and two other partly overlapping states (sensorimotor and DMN). This finding is important because the feedback never conveyed specific information about transition probabilities, only about occurrence of the PH-state. Yet, this was sufficient to help participants engage in the brain mechanism underlying riPH, changes in transitions favoring the PH-state ³⁵, as well as facilitation towards the sensorimotor and DMN states, which were similar to the PH-state. Studying the topologies of these three brain states, we note that overlap occurred mostly in pSTS, IPL, and dorso-lateral prefrontal-cortex. The former two regions have been related to riPH ^{33,35}, clinical PH ²⁷, and are cross-modal multisensory integration hubs ⁴⁹⁻⁵¹. The IPL has also been linked to cognitive and attentional control ⁵¹ and the dlPFC to cognitive control related to NF ^{36,52}. While not object of direct study in the present work, we hypothesize that the shared regions between brain states might have diminished the energy ⁵³ required to transition between the states ^{54,55}, guaranteeing an ease of transitions back to the PH-state after leaving it. We propose that facilitation of transitions to specific brain states might constitute a more mechanism underlying hallucinations in general, that ultimately participants were able to develop through our fMRI-NF protocol, because changes in transition probabilities observed in the current NF work and previous riPH work ³⁵, are also supported by clinical studies unrelated to NF. In schizophrenia, for example reduced dwell time in specific brain states is related to hallucination proneness ⁵⁶, and specifically for transition between brain states, work in PD identified that aberrant facilitation of transitions to specific brain states marks PD patients suffering from visual hallucinations ⁵⁵. In sum, we identified that control over the NF-signal was achieved by modulating specific transitions between different brain states, and showed for the first time that NF regulation can be achieved through changes in transition probabilities across brain states.

In addition to up-regulation, participants also learned to down-regulate the PH-state, but during the control-NF condition rather than during PH-NF. Down-regulation was characterized by a decreased

Part I: Identifying and modulating neural correlates of PH

recruitment of the PH-state and higher recruitment of the sensorimotor+ state, instead. Given that the latter contained deactivations of the PH-state, the avoidance of the PH-state was also accompanied by its partial deactivation. This deactivation mechanism was not simply a compensation for hyper-activation of the PH-state during up-regulation, because hyper-activation was not observed, and because a compensation mechanism would have to be present also in rest following up-regulation (again not observed in the present study).

The development of these NF strategies led to a specific increase in riPH after NF training, but not of other mental states that typically accompany riPH (passivity experiences, loss of agency, or control sensations), despite participants being fully blinded to the goals of the NF training. This provides evidence of the neural-specificity between the targeted fronto-parietal PH-network and our behavioural measure, and provides first causal evidence of the role of this network in PH. This finding was already robust, but we hypothesize its effect could have been increased, had our participants relied solely on the up-regulation strategy.

The development of these NF strategies led to a specific increase in riPH after NF training, but not of any other mental state that typically accompanies riPH (passivity experiences, loss of agency, nor control sensations), even though participants were fully blinded to the NF-training goals. While this finding is robust, we hypothesize its effect could have been increased, had our participants relied solely on the up-regulation strategy. These data were further corroborated by the identification of sustained neural changes post-training, similar to the NF regulation mechanisms, in participants that achieved most HRP runs and became sensitive to riPH post-training. On one hand, up-regulation observed during PH-NF runs, was followed in the phase post-training by higher occurrences of the PH-state during the PH-inducing condition. On the other hand, down-regulation observed during control-NF, was followed post-training by decreased occurrences of the PH-state during the control condition. Based on these findings, we propose that the sustained higher occurrences during the PH-inducing condition facilitated participants to enter in the transition probability pattern that characterizes riPH³⁵, whereas the sustained decrease in occurrences made it harder to enter such a state during the control condition.

Several NF studies have achieved modulation over different cognitive or behavioural aspects, such as attention^{39,59}, fear⁶⁰, and working memory⁵². Of particular interest for the current study are comparisons to other NF studies targeting hallucinations. One of the first NF studies to do so, asked 12 schizophrenic patients with auditory-verbal hallucinations (AVH) to decrease the activity of the posterior superior temporal gyrus (pSTG)⁵⁷. While the study was the first showing that these patients could down-regulate pSTG activity, this was not associated with any significant change in AVH severity (only a non-targeted increase in FC between pSTS and IFG (amongst other increases in FC), significantly correlated with a decrease in AVH). Moreover, as no control condition was performed, a confound of placebo/intervention effect remained in this seminal work⁴⁸. Another NF study directly targeted down-regulation of pSTG-IFG connectivity, in patients with schizophrenia suffering from AVH⁵⁸. Although a modulation of the targeted connectivity was achieved, this again did not correlate with decreased AVH severity. Again, only a non-targeted connectivity between pSTG, IFG, and IPL was associated with subjective rating of symptom severity (which was not specific to AVH) and no relation with the burden of positive symptoms was found. While these two studies showed that patients with schizophrenia can modulate pSTS-related FC, no specific changes were found between NF regions and AVH reduction. Finally, a more recent study conducted a single session of explicit NF training (i.e., using an enforced meditation strategy) targeting anti-correlations between the DMN and an executive

Part I: Identifying and modulating neural correlates of PH

network in order to reduce AVH in patients suffering from schizophrenia ⁴³. Although, an increase in anti-correlation between these networks correlated with a decrease of AVH burden after NF training, the control condition included in this study, which was NF-training of motor regions during finger-tapping, did not appropriately control for the major confound of the explicit meditation strategy used, especially when considering that meditation alone was shown to reduce hallucination severity ⁶¹.

One important difference compared to this previous work, is that in our NF study we targeted mechanisms of a specific hallucination while it was induced by our robotic method. This, and the fact that it was done in the general population, relieved us of the confound of disease trait and guaranteed we targeted a hallucination state. We further ensured our paradigm allowed bi-directional regulation, which best controls for various NF regulation confounds ⁴⁸. By separately analysing the regulation mechanisms and pre/post training neural changes, we guaranteed independence between the two, but still recovered sustained neural changes that matched the identified regulation mechanisms. Finally, while our participants were blinded to the goals of the NF task, the changes in behaviour post-training led to changes in the sensitivity of riPH in the fifth session, with control analyses for neural changes related to other sensations revealing no changes post-training. Together our results demonstrate considerable neurospecificity between the PH-state and sensitivity to riPH, and furthermore represent the first time that any study has successfully achieved modulation of hallucination in such a specific manner paired with sustained changes. They are also the first time, NF is shown to modulate transitions and occurrences of brain states.

The present approach may have translational potential and may allow to modulate and decrease hallucination severity. There are several limitations of the present approach that should be pointed out. While it was possible to confirm sustained neural changes after training through both successful NF regulation and an increase in sensitivity to riPH post-training, it is unclear whether re-instantiation, avoidance or a combination of both strategies, modulated riPH sensitivity. Additionally, in the current experiment the number of participants that were sensitivity to riPH at the start was lower than what we have observed in other experiments both in the MRI ³³ and outside ^{27,28,32}. It is possible that this benefitted the effect size of NF-training on sensitivity to riPH, for example by accidental sampling of participants with low sensitivity. However, it could be argued that avoiding ceiling effects is actually desirable in the current experiment. Furthermore, any spurious general effect was controlled by analyzing various other sensations whose underlying neural mechanisms were not targeted, and which showed no effect of NF-training.

Our goal was to assess whether volitional control over a brain state related to PH-induction was possible, and whether such volitional control would change sensitivity to riPH. Our findings confirm that volitional control is achievable, modifies sensitivity to the associated hallucination, and leads to sustained neural changes in successful and sensitive participants. Considering that hallucinations are major symptoms in several frequent diseases ^{6,8} and can serve as markers for disease progression ^{8,24}, we propose that the relevance of this study goes beyond the specific case of PH and neuroscience of hallucinations. Supported by evidence that personal coping strategies can prevent the progression of neuropsychiatric symptoms in PD if done early ⁶², we propose that NF-training paired with our robotic setup, which has been shown to trigger PH in these patients ³³, could be used as a means to decrease the activity of the hallucination brain state and potentially slowdown or even stop the progression of neuropsychiatric and neurocognitive dysfunction in PD patients.

2.2.5 Methods

2.2.5.1 Participants

28 healthy individuals were recruited to take part in this experiment. Out of these, two participants retracted from the experiment voluntarily citing time constraints, one participant became unreachable throughout the sessions, another one was excluded due to the use of psychoactive substances between sessions, and four did not pass the anamnestic interview (below) due to substance use, a neurological atrophy disorder, and potential psychiatric disorders. The remaining 20 participants (13 male) had an average age of 25.25 years (± 3.19 ; range: 21-32), and were all right-handed according to the Edinburgh Handedness Inventory 63 (0.68 ± 0.19 ; range: 0.35 – 1). All participants gave their informed consent prior to the start of the experiment, following the Declaration of Helsinki, and the study was approved by the local ethics committee of the Canton of Geneva, Switzerland.

2.2.5.2 Interview

Prior to the start of the experiment, participants took part in a screening session, to evaluate their current medical status, use of medications affecting the nervous and muscular system, history of neurological disorders, substance abuse, and psychiatric disorders. The full interview can be found in Annex 1. Recruited participants did not have any diagnosed neurological or psychiatric disorder, nor used medications or psychoactive drug for the period of two weeks preceding the start of the experiment. Recruited participants were routinely asked about their use of medications, and recreational drugs.

2.2.5.3 MRI Data Acquisition

Functional image acquisitions were performed at the MRI facility of the Campus Biotech (Geneva, Switzerland), on a Siemens MAGNETOM Prisma 3T scanner, using a 64-channel head-and-neck coil. For the resting-state, sensorimotor task, and NF task, echo-planar sequences were used (EPI, TR = 1.5s, TE = 31 ms, with a flip angle of 64°, GRAPPA = 4), with an in-plane resolution of 2.0x2.0 mm² and slice thickness of 2.0 mm (no gap, 68 slices FOV = 216 mm). Anatomical images were acquired with a T1-weighted MPAGE sequence (192 slices, FOV = 256 mm, TR = 2.2s, TE = 2.96ms).

2.2.5.4 Robotic Procedure for PH-induction

The tactile stimulation presented to the participants during the experiment was administered by a robotic system composed by two main components, a front-robot, and a back-robot concealed below an adapted platform-bed in the MRI. The front robot included an extended lever that participants used to manipulate the front-robot by performing movements with their right arm. Participants were instructed to pivot movements from the elbow and wrist,

but to avoid moving the upper arm and shoulder as to minimise head movements. The back-robot then delivered tactile feedback to the participant's back (through a gap in the middle of the MRI platform bed) by mimicking the participant's movements either in real-time (synchronous condition – non inducing), or with a small delay (500 milliseconds – asynchronous condition – PH-inducing).

2.2.5.5 Experimental Paradigm

Participants completed five fMRI sessions on five different days, with at least one day of break between each session. A significant effort was done to have all participants complete the experiment under two weeks, however, due to the COVID-19 pandemic this was not possible for one participant who completed it in 28 days, and another who attended day 1 twenty days before starting day 2 (first day of training), and then completed the rest in 12 days. The average completion time excluding these two cases was 12.52 days (± 3.66 ; range: 8-19 days).

The first day was dedicated to the assessment of the participants' baseline sensitivity to PH induction before NF training. The second, third, and fourth days, were dedicated to NF training and transfer runs. In particular, we aimed at 3 runs of NF per day followed by one transfer run. The fifth day was identical to the first, dedicated to the assessment of sensitivity to PH induction after NF training. For all parts of the study, participants had their eyes covered as typically done in previous PH-inducing paradigms³², and wore in-ear earphones that were used to both talk with the participants during breaks and to provide auditory feedback during NF runs.

Days 1 and 5: Assessment of PH induction Participants performed two runs of the robot manipulation task, devised to induce PH^{32,34} in the fMRI scanner. Each run consisted of 16 blocks of 30 seconds of robot manipulation interleaved by periods of rest of 15 seconds. Robot manipulation was performed either in the PH-inducing (asynchronous manipulation), or in the non-inducing (synchronous manipulation) conditions. Next, participants performed two isolated blocks of robot manipulation in each condition (random order, counterbalanced across participants), each followed by a questionnaire tailored to assess the intensity of different subjective experiences during each condition (7-point Likert scale, Supplementary Table 27). This included but was not limited to: PH - "I felt as if someone was behind me", Passivity Experiences – "I felt as if someone else was touching my body", and Loss of Agency – "I felt as if I was not in control of my movements or actions".

Days 2, 3, and 4: NF Training During the NF training sessions, participants performed three runs of NF, followed by one transfer run. The NF runs were composed by one habituation block, and six NF blocks. The habituation block was identical to the NF blocks (further described); however, no feedback was given. This block served to establish a baseline for each session. NF blocks were each made up of a rest condition, followed by a control condition, and an up-regulation condition. During the rest condition, participants laid in the scanner without

moving the robotic system, and without receiving any feedback, for 14 seconds. After a single beep, the control condition started, and would last for a duration of 21 seconds. Here, participants started moving the front-robot, in the synchronous condition, meaning they received tactile stimulation to their back from the back-robot in real-time with their movements. Participants also listened to the minimum level of auditory feedback during this condition, regardless of ongoing brain activity, so that control and up-regulation were matched as closely as possible. Furthermore, participants were instructed that at this point they should not engage in any regulation strategy, and instead should try to “clear their minds”. After two beeps, the up-regulation condition started for a duration of 60 seconds. Here, the robot switched to the asynchronous condition, providing tactile feedback to the participant’s back with a delay of 500 milliseconds with respect to their own movements. The auditory feedback became variable and derived from the participants’ ongoing brain activity. This feedback was continuous and provided as closely to real-time as possible, meaning with a delay of at least the haemodynamic delay (~5 seconds) and processing time (~1 second). Participants were made aware of the feedback’s intrinsic delay. They were also instructed that during this condition, they should try to achieve the highest level of auditory feedback (see next sub-chapter) by means of a mental strategy of their choice. More details regarding the instructions to participants can be found on “Guidance for participants regarding NF training” on the supplemental information.

Transfer runs had an identical block design to the NF runs, however, there were only two blocks during which no feedback was presented. Participants were informed they should still try to deploy the mental strategy learned during the up-regulation condition.

2.2.5.6 Real-time data processing and neurofeedback

For real-time processing of incoming functional images, and presentation of feedback, we made use of the open-source software OpenNFT⁶⁴, running on MATLAB 2019b and Python 3.5. In-house adaptations were made to provide a feedback score based on full-brain correlations, to provide auditory rather than visual feedback, and to have OpenNFT communicate with another computer in the network which was controlling the robotic device used for PH induction.

Pre-processing Throughout the runs, data was directly reconstructed in the scanner’s console and sent to the computer running OpenNFT. The standard OpenNFT pre-processing pipeline was used, which included spatial realignment to the template scan (obtained in a short sequence at the beginning of each day), re-slicing to compensate for head movements, and spatial smoothing (FWHM = 5 mm).

Processing Linear trends were removed from each voxel in real-time, using an adapted filter which also functioned as a high-pass filter. Before being used for NF computations, the volumes were also masked with a participant-specific grey matter mask.

Neurofeedback The processed and detrended data of each full-brain volume was directly correlated with the PH-network brain map as obtained from a previous independent dataset³⁶. This spatial correlation between ongoing brain activity and PH network (NF-signal) was then

Part I: Identifying and modulating neural correlates of PH

used to compute a score that was fed back to the participants in auditory form as one of eight levels.

During up-regulation, the feedback score at each time point, was computed as the NF-signal, minus the median NF-signal of the previous control condition, normalised by the minimum and maximum NF-signal values obtained in the window with a length of one block (67 scans, i.e. dynamic range) (eq.1). Crucially, this guaranteed that the feedback during the up-regulation condition had an adaptive baseline, as assessed by the previous control condition, and could account for movement performed with the robot.

$$\text{Eq. 1} \quad \text{Feedback Score} = \frac{\text{NFsignal} - \text{median}(\text{NFsignal})_{\text{control}}}{\text{max}(\text{NFsignal})_{\text{DR}} - \text{min}(\text{NFsignal})_{\text{DR}}}$$

Importantly, the dynamic range was also continuously adapted according to each participant's performance. At the transition point from the control to the up-regulation condition, and if a participant's feedback-score was in the upper quartile of the dynamic range for more than 80% of the previous up-regulation condition, then the upper limit was increased by 10% to make the task slightly more challenging. Conversely, if the participant's feedback-score was in the upper quartile of the dynamic range for less than 60% of the previous up-regulation condition, then the upper limit was decreased by 10% to ease the task. No changes were applied if the feedback-score was in the upper quartile for 60 to 80% of the previous up-regulation condition. This process helped keep engagement for both the participants that were very good at the task, and for participants that had more difficulties. The reference values were determined during pilot experiments.

The auditory feedback consisted of an amplitude-modulated sound with main harmonic frequency of 50Hz (complemented by second and third harmonics), producing a sound comparable to that of a wave approaching and receding. At its minimum level, the auditory feedback had an amplitude modulation following a frequency that allowed two waves per TR (~1.33Hz). The number of waves heard per TR gradually increased by two, over the next six levels (level 2 had four waves per TR, ~2.66Hz, level 3 had six waves per TR, 4Hz, and so on). At the highest level (level 8), no frequency modulation was used allowing the reproduction of a constant tone. This allowed for participants to always implicitly know the highest level of auditory feedback, which was when the sound became constant (as explained to the participants: when the waves in the sound become "so fast" that they disappear). Simultaneously, participants were also constantly aware of the minimum level of feedback since this was presented during the previous control condition. This design guaranteed the typical advantages of a visual feedback (i.e., constant awareness of the minimum and maximum of the feedback on the display), but in an implicit manner. We nonetheless guaranteed that every training day, before the experiment begun, but already inside the MRI, participant's re-heard all auditory feedback levels.

Participants were informed that the auditory feedback was related to their own brain activity, however, they were blinded as to which specific aspect of brain activity the NF was related to (i.e., they did not know the measured brain activity was related to the occurrence of a network related to PH induction). Participants were told that they should engage in mental strategies during the up-regulation condition, but not during the control condition. They were told that

their regulation performance could benefit from trying to “clear their minds” during the control condition. Further instructions regarding strategies were given: no engagement in physical strategies such as moving parts of their body unrelated to robot manipulation; no explicit change of the pace of the robot manipulation between up-regulation and control; no engagement in breathing strategies; no strategies based on the sounds produced by the MR scanner.

2.2.5.7 Data Analysis (offline – after all acquisitions)

Neurofeedback Training Performance To assess the performance of each participant in NF regulation, we used non-parametric hypothesis testing. The null hypothesis is operationalized using surrogate data generation. First, for each NF run we generate 100.000 surrogates of the original NF-signal (figure 2A) by phase randomization of the Fourier coefficients of the time series 65 (figure 2B). Importantly, the surrogate time series obtained in this way preserve the autocorrelation properties of the empirical one. Second, we compute the average feedback score for each surrogate, generating a null distribution of mean up-regulation feedback-scores for the surrogates (figure 2B). Third, we establish a percentile threshold above which a run is considered as having high-regulation performance (HRP; figure 2B). Given the adaptive nature of the feedback, high percentile thresholds selected runs that showed more evidence for modulation of the PH-network between up-regulation and control conditions, while accounting for the NF-signal variability across runs and across participants’ task performance. While this was a selection threshold, we then used a Binomial distribution to compute the probability of observing that many successes (at the group level) as compared to chance – with chance here being defined by the threshold (i.e. a threshold of 70% would define random chance as 70% non-HRP and 30% HRP). The percentile threshold was chosen at 84.1%. This decision was made as to prune runs with higher modulation between up-regulation and control conditions, but importantly, we demonstrate that the probability of obtaining the real feedback scores under the null distribution hypothesis of the surrogates is always extremely low for any percentile threshold chosen (figure 2D).

MRI Data Preprocessing Offline data preprocessing was done with standard CONN pipelines 66, that in turn make additional use of SPM (Wellcome Department of Cognitive Neurology, Institute of Neurology, UCL, London, UK) and ART toolbox functionalities (Gabrielli lab, MIT). This pipeline included, spatial re-alignment and re-slicing, identification and removal of high-movement frames, followed by tri-linear interpolation of those data-points. Normalisation to MNI space (resolution), and spatial smoothing (FMWH) followed. Data denoising was further performed with the COON toolbox to remove potential confounds related to, respiration, movement, and related to the previous scrubbing-interpolation procedure. Linear trends were also removed.

Brain State Extraction through Co-Activation Pattern Analysis We performed CAP analysis to extract the main brain states active during NF runs (days 2, 3, and 4), transfer runs (days 2, 3, and 4), and PH induction runs (days 1, and 5). This was a two-step procedure.

For the NF runs, we followed the same procedure used in previous work, that originally identified the PH-network 36. In brief, we extracted brain states that co-fluctuated with the activity in two seed regions relevant for the induction of PH 34, the posterior superior

Part I: Identifying and modulating neural correlates of PH

temporal sulcus (pSTS), and the inferior frontal gurus (IFG), using the TbCAPs toolbox 67. CAP analysis performs clustering with a k-means algorithm, of selected timepoints of the acquisition, where one or both of the seed regions exceed a z-score of 1 (time-points with a framewise displacement of 0.5 mm were scrubbed). The number of k states for clustering was set at eight based on consensus clustering 68, a measure of clustering stability across clustering's performed with different number of k states (Supplemental Figs 1 and 2). Finally, representative states were obtained by averaging the time-frames that are attributed to the same k state. Timepoints were none of the seeds exceeded a z-score of 1 in activity, were marked as inactive state (state0).

We used a spatial similarity procedure integrated in the TbCAPs, to attribute the CAPs identified during the NF runs, to the PH induction runs and to the transfer runs. In sum, first the correlation between a scan and each of the states is computed, second, a state is assigned to that scan, if the correlation value is above the 5th percentile of the distribution of correlations for that state.

State/CAP metrics We investigated the dynamics of brain states through two metrics: occurrences, and transition probabilities. Occurrences can be defined as the perceptual occurrence of a state in the entire length of a condition. Hence, if the set $D_{S,C}$ contains all timepoints of a condition C for a participant S, the occurrences of state i are defined as the sum of an indicator function $State_i[k]$ over the timepoints k , divided by the length of the set $D_{S,C}$ (eq. 3). Transition probabilities, refer to how different states transition amongst themselves in time, and gives the probabilities of having a certain next state (j) after an initial state (i) (eq. 5). This should be seen as a 1st order Markov chain.

$$\text{Eq. 2} \quad Occ(i; D_{S,C}) = \frac{\sum_{k \in D_{S,C}} State_i[k]}{|D_{S,C}|}.$$

$$\text{Eq. 3} \quad TP(i, j; D_S) = \frac{\sum_{k \in D_S} State_i[k] State_j[k+1]}{\sum_{k \in D_S} State_i[k]}$$

State/CAP metrics modelling – Neurofeedback and Transfer runs To identify how brain dynamics are modulated by the different tasks, we have modelled the above state metrics using Linear Mixed Models (LMM) using the package lme4⁶⁹ available for R (version 4.0.0).

For the NF and transfer runs, we modelled the metrics of occurrences of the states, with predictors for state, condition, number of the run (absolute over days), and a binary predictor defining if the run was HRP or not. Included in the model were also random-effect predictors for day and participants (example for modelling occurrences: eq. 4).

$$\text{Eq. 4} \quad Occ \sim State * HRP * Condition * Run\ Number + (1|Day) + (1|Participant)$$

Part I: Identifying and modulating neural correlates of PH

Model comparison based on the Akaike Information Criteria ⁷⁰ (AIC) revealed that the predictor for number of the run, gravely increased the complexity of the model, without adding explained variance. This predictor was hence dropped from the model.

Transition probabilities for the CAPs, were modelled in the same manner, however, instead of a predictor for state being used, two other predictors, for initial state and next state were used (see example of eq. 5).

Notably our analysis strategy started with the highest hierarchical model (eq. 4), and if an interaction was detected we would then seek to further investigate that effect by decomposing the model. However, although our experiment included a considerable amount of data, the number of multiple comparisons for both these occurrences and transition probabilities can grow very fast due to the number of levels of the predictors. To address this we used an analysis strategy where models for example with 3-way interactions were decomposed for only one factor (rather than two), and further decomposition would only proceed for intermediate models where the remaining 2-way interaction was maintained. For example, if a model showed a triple interaction with state, HRP and, condition, this model would be first decomposed for condition, which only has 3 levels, rather than immediately decomposed for all levels of conditions and state (30 levels). Analysis would then only proceed with decomposition per states, if an interaction state:condition was maintained in the intermediate model.

State/CAP metrics modelling – PH induction task To identify how brain dynamics changed due to the NF training targeting the PH network, we modelled the difference in state metrics between day 5 and day 1. For occurrences, this was done using predictors for state, condition, number of high-regulation performance runs during training, and sensitivity to PH after training (day 5).

For transition probabilities, instead of the single predictor for state, two predictors for initial state and next state were used, as seen in eq. 5.

$$\text{Eq. 5 } TP \sim \text{State}_{init} * \text{State}_{next} * \text{Number of HPRs} * \text{Condition} * \text{PH sensitivity}_{day5} + (1|Day) + (1|Participant)$$

2.2.6 References

1. Corlett, P. R. et al. Hallucinations and Strong Priors. *Trends in Cognitive Sciences* 23, 114–127 (2019).
2. Larøi, F. et al. An epidemiological study on the prevalence of hallucinations in a general-population sample: Effects of age and sensory modality. *Psychiatry Research* 272, 707–714 (2019).
3. Badcock, J. C. et al. Hallucinations in Older Adults: A Practical Review. *Schizophrenia Bulletin* 46, 1382–1395 (2020).
4. Badcock, J. C., Dehon, H. & Larøi, F. Hallucinations in Healthy Older Adults: An Overview of the Literature and Perspectives for Future Research. *Front. Psychol.* 8, 1134 (2017).
5. Hayes, J. & Leudar, I. Experiences of continued presence: On the practical consequences of ‘hallucinations’ in bereavement. *Psychol Psychother Theory Res Pract* 89, 194–210 (2016).
6. Millan, M. J. et al. Altering the course of schizophrenia: progress and perspectives. *Nat Rev Drug Discov* 15, 485–515 (2016).
7. McKeith, I. G. et al. Diagnosis and management of dementia with Lewy bodies: Fourth consensus report of the DLB Consortium. *Neurology* 89, 88–100 (2017).
8. ffytche, D. H. et al. The psychosis spectrum in Parkinson disease. *Nat Rev Neurol* 13, 81–95 (2017).
9. Diederich, N. J., Fénelon, G., Stebbins, G. & Goetz, C. G. Hallucinations in Parkinson disease. *Nat Rev Neurol* 5, 331–342 (2009).
10. Wackermann, J., Putz, P. & Allefeld, C. Ganzfeld-induced hallucinatory experience, its phenomenology and cerebral electrophysiology. *Cortex* 44, 1364–1378 (2008).
11. Allefeld, C., Pütz, P., Kastner, K. & Wackermann, J. Flicker-light induced visual phenomena: Frequency dependence and specificity of whole percepts and percept features. *Consciousness and Cognition* 20, 1344–1362 (2011).
12. Baggott, M. J. et al. Investigating the Mechanisms of Hallucinogen-Induced Visions Using 3,4-Methylenedioxyamphetamine (MDA): A Randomized Controlled Trial in Humans. *PLoS ONE* 5, e14074 (2010).
13. Timmermann, C. et al. DMT Models the Near-Death Experience. *Front. Psychol.* 9, 1424 (2018).
14. Peter Brugger, Marianne Regard, & Theodor Landis. Brugger et al 1996 - Uniilaterally felt presences The neuropsychiatry of ones invisible doppelganger.pdf. *Neuropsychiatry Neuropsychology and Behavioral Neurology* 2, 19–38 (1996).

Part I: Identifying and modulating neural correlates of PH

15. Llorca, P. M. et al. Hallucinations in schizophrenia and Parkinson's disease: an analysis of sensory modalities involved and the repercussion on patients. *Sci Rep* 6, 38152 (2016).
16. Nagahama, Y. et al. Classification of Psychotic Symptoms in Dementia With Lewy Bodies. *The American Journal of Geriatric Psychiatry* 15, 961–967 (2007).
17. Nicastro, N., Eger, A. F., Assal, F. & Garibotto, V. Feeling of presence in dementia with Lewy bodies is related to reduced left frontoparietal metabolism. *Brain Imaging and Behavior* 14, 1199–1207 (2020).
18. Fenelon, G., Soulas, T., de Langavant, L. C., Trinkler, I. & Bachoud-Levi, A.-C. Feeling of presence in Parkinson's disease. *Journal of Neurology, Neurosurgery & Psychiatry* 82, 1219–1224 (2011).
19. Blanke, O., Ortigue, S., Coeytaux, A., Martory, M.-D. & Landis, T. Hearing of a Presence. *Neurocase* 9, 329–339 (2003).
20. Arzy, S., Seeck, M., Ortigue, S., Spinelli, L. & Blanke, O. Induction of an illusory shadow person. *Nature* 443, 287–287 (2006).
21. Peter Brugger, Marianne Regard, Theodor Landis, & O Oelz. Brugger et al 1999 - Hallucinatory experiences in extreme-altitude climbers.pdf. *Neuropsychiatry Neuropsychology and Behavioral Neurology* 12, 67–71 (1999).
22. Pagonabarraga, J. et al. Minor hallucinations occur in drug-naïve Parkinson's disease patients, even from the premotor phase: Minor Hallucinations in Untreated PD Patients. *Mov Disord.* 31, 45–52 (2016).
23. Goetz, C. G., Fan, W., Leurgans, S., Bernard, B. & Stebbins, G. T. The Malignant Course of "Benign Hallucinations" in Parkinson Disease. *Arch Neurol* 63, 713 (2006).
24. Lenka, A., Pagonabarraga, J., Pal, P. K., Bejr-Kasem, H. & Kulisvesky, J. Minor hallucinations in Parkinson disease: A subtle symptom with major clinical implications. *Neurology* 93, 259–266 (2019).
25. Hely, M. A., Reid, W. G. J., Adena, M. A., Halliday, G. M. & Morris, J. G. L. The Sydney multicenter study of Parkinson's disease: The inevitability of dementia at 20 years: Twenty Year Sydney Parkinson's Study. *Mov. Disord.* 23, 837–844 (2008).
26. Ravina, B. et al. Diagnostic criteria for psychosis in Parkinson's disease: Report of an NINDS, NIMH work group. *Mov Disord.* 22, 1061–1068 (2007).
27. Blanke, O. et al. Neurological and Robot-Controlled Induction of an Apparition. *Current Biology* 24, 2681–2686 (2014).
28. Serino, A. et al. Thought consciousness and source monitoring depend on robotically controlled sensorimotor conflicts and illusory states. *iScience* 24, 101955 (2021).
29. Orepic, P., Rognini, G., Kannape, O. A., Faivre, N. & Blanke, O. Sensorimotor conflicts induce somatic passivity and louden quiet voices in healthy listeners. *Schizophrenia Research* 231, 170–177 (2021).

Part I: Identifying and modulating neural correlates of PH

30. Faivre, N. et al. Sensorimotor conflicts alter metacognitive and action monitoring. *Cortex* 124, 224–234 (2020).
31. Bernasconi, F. et al. Neuroscience robotics for controlled induction and real-time assessment of hallucinations. *Nat Protoc* (2022) doi:10.1038/s41596-022-00737-z.
32. Salomon, R. et al. Sensorimotor Induction of Auditory Misattribution in Early Psychosis. *Schizophrenia Bulletin* 46, 947–954 (2020).
33. Bernasconi, F. et al. Robot-induced hallucinations in Parkinson’s disease depend on altered sensorimotor processing in fronto-temporal network. *Sci. Transl. Med.* 13, eabc8362 (2021).
34. Hara, M. et al. A novel manipulation method of human body ownership using an fMRI-compatible master–slave system. *Journal of Neuroscience Methods* 235, 25–34 (2014).
35. Dhanis, H. et al. Robotically-induced hallucination triggers subtle changes in brain network transitions. *NeuroImage* 248, 118862 (2022).
36. Sitaram, R. et al. Closed-loop brain training: the science of neurofeedback. *Nat Rev Neurosci* 18, 86–100 (2017).
37. Watanabe, T., Sasaki, Y., Shibata, K. & Kawato, M. Advances in fMRI Real-Time Neurofeedback. *Trends in Cognitive Sciences* 21, 997–1010 (2017).
38. Thibault, R. T., MacPherson, A., Lifshitz, M., Roth, R. R. & Raz, A. Neurofeedback with fMRI: A critical systematic review. *NeuroImage* 172, 786–807 (2018).
39. Pamplona, G. S. P. et al. Network-based fMRI-neurofeedback training of sustained attention. *NeuroImage* 221, 117194 (2020).
40. Krause, F. et al. Self-regulation of stress-related large-scale brain network balance using real-time fMRI neurofeedback. *NeuroImage* 243, 118527 (2021).
41. Pindi, P., Houenou, J., Piguet, C. & Favre, P. Real-time fMRI neurofeedback as a new treatment for psychiatric disorders: A meta-analysis. *Progress in Neuro-Psychopharmacology and Biological Psychiatry* 119, 110605 (2022).
42. Yamada, T. et al. Resting-State Functional Connectivity-Based Biomarkers and Functional MRI-Based Neurofeedback for Psychiatric Disorders: A Challenge for Developing Theranostic Biomarkers. *International Journal of Neuropsychopharmacology* 20, 769–781 (2017).
43. Bauer, C. C. C. et al. Real-time fMRI neurofeedback reduces auditory hallucinations and modulates resting state connectivity of involved brain regions: Part 2: Default mode network -preliminary evidence. *Psychiatry Research* 284, 112770 (2020).
44. Liu, X. & Duyn, J. H. Time-varying functional network information extracted from brief instances of spontaneous brain activity. *Proc. Natl. Acad. Sci. U.S.A.* 110, 4392–4397 (2013).

Part I: Identifying and modulating neural correlates of PH

45. Raichle, M. E. et al. A default mode of brain function. *Proc. Natl. Acad. Sci. U.S.A.* 98, 676–682 (2001).
46. deCharms, R. C. et al. Control over brain activation and pain learned by using real-time functional MRI. *Proc. Natl. Acad. Sci. U.S.A.* 102, 18626–18631 (2005).
47. Subramanian, L. et al. Real-Time Functional Magnetic Resonance Imaging Neurofeedback for Treatment of Parkinson’s Disease. *Journal of Neuroscience* 31, 16309–16317 (2011).
48. Sorger, B., Scharnowski, F., Linden, D. E. J., Hampson, M. & Young, K. D. Control freaks: Towards optimal selection of control conditions for fMRI neurofeedback studies. *NeuroImage* 186, 256–265 (2019).
49. Beauchamp, M. S. See me, hear me, touch me: multisensory integration in lateral occipital-temporal cortex. *Current Opinion in Neurobiology* 15, 145–153 (2005).
50. Beauchamp, M. S., Yasar, N. E., Frye, R. E. & Ro, T. Touch, sound and vision in human superior temporal sulcus. *NeuroImage* 41, 1011–1020 (2008).
51. Seghier, M. L. The Angular Gyrus: Multiple Functions and Multiple Subdivisions. *Neuroscientist* 19, 43–61 (2013).
52. Sherwood, M. S., Kane, J. H., Weisend, M. P. & Parker, J. G. Enhanced control of dorsolateral prefrontal cortex neurophysiology with real-time functional magnetic resonance imaging (rt-fMRI) neurofeedback training and working memory practice. *NeuroImage* 124, 214–223 (2016).
53. Adams, R. A., Stephan, K. E., Brown, H. R., Frith, C. D. & Friston, K. J. The Computational Anatomy of Psychosis. *Front. Psychiatry* 4, (2013).
54. Betzel, R. F., Gu, S., Medaglia, J. D., Pasqualetti, F. & Bassett, D. S. Optimally controlling the human connectome: the role of network topology. *Sci Rep* 6, 30770 (2016).
55. Zarkali, A. et al. Changes in dynamic transitions between integrated and segregated states underlie visual hallucinations in Parkinson’s disease. *Commun Biol* 5, 928 (2022).
56. Weber, S. et al. Dynamic Functional Connectivity Patterns in Schizophrenia and the Relationship With Hallucinations. *Front. Psychiatry* 11, 227 (2020).
57. Orlov, N. D. et al. Real-time fMRI neurofeedback to down-regulate superior temporal gyrus activity in patients with schizophrenia and auditory hallucinations: a proof-of-concept study. *Transl Psychiatry* 8, 46 (2018).
58. Zweerings, J. et al. Neurofeedback of core language network nodes modulates connectivity with the default-mode network: A double-blind fMRI neurofeedback study on auditory verbal hallucinations. *NeuroImage* 189, 533–542 (2019).
59. deBettencourt, M. T., Cohen, J. D., Lee, R. F., Norman, K. A. & Turk-Browne, N. B. Closed-loop training of attention with real-time brain imaging. *Nat Neurosci* 18, 470–475 (2015).

Part I: Identifying and modulating neural correlates of PH

60. Koizumi, A. et al. Fear reduction without fear through reinforcement of neural activity that bypasses conscious exposure. *Nat Hum Behav* 1, 0006 (2017).
61. Sheng, J., Yan, Y., Yang, X., Yuan, T. & Cui, D. The effects of Mindfulness Meditation on hallucination and delusion in severe schizophrenia patients with more than 20 years' medical history. *CNS Neurosci Ther* 25, 147–150 (2019).
62. Diederich, N. J., Pieri, V. & Goetz, C. G. Coping strategies for visual hallucinations in Parkinson's disease. *Mov Disord.* 18, 831–832 (2003).
63. R. C. Oldfield. Oldfield 1971 - The assessment and analysis of handedness The Edinburgh Inventory.pdf. *Neuropsychologia* 9, 97–113 (1971).
64. Koush, Y. et al. OpenNFT: An open-source Python/Matlab framework for real-time fMRI neurofeedback training based on activity, connectivity and multivariate pattern analysis. *NeuroImage* 156, 489–503 (2017).
65. Theiler, J., Eubank, S., Longtin, A., Galdrikian, B. & Doyne Farmer, J. Testing for nonlinearity in time series: the method of surrogate data. *Physica D: Nonlinear Phenomena* 58, 77–94 (1992).
66. Whitfield-Gabrieli, S. & Nieto-Castanon, A. Conn : A Functional Connectivity Toolbox for Correlated and Anticorrelated Brain Networks. *Brain Connectivity* 2, 125–141 (2012).
67. Bolton, T. A. W. et al. TbCAPs: A toolbox for co-activation pattern analysis. *NeuroImage* 211, 116621 (2020).
68. Stefano Monti, Pablo Tamayo, Jill Mesirov, & Todd Golub. Consensus Clustering: A Resampling-Based Method for Class Discovery and Visualization of Gene Expression Microarray Data. *Machine Learning* 52, 91–118 (2003).
69. Bates, D., Mächler, M., Bolker, B. & Walker, S. Fitting Linear Mixed-Effects Models Using lme4. *J. Stat. Soft.* 67, (2015).
70. Hirotugu Akaike. Akaike 1974 - A new look at the statistical model identification.pdf. *IEEE Transactions on automatic control* 19, 716–723 (1974).

2.2.7 Extended Data

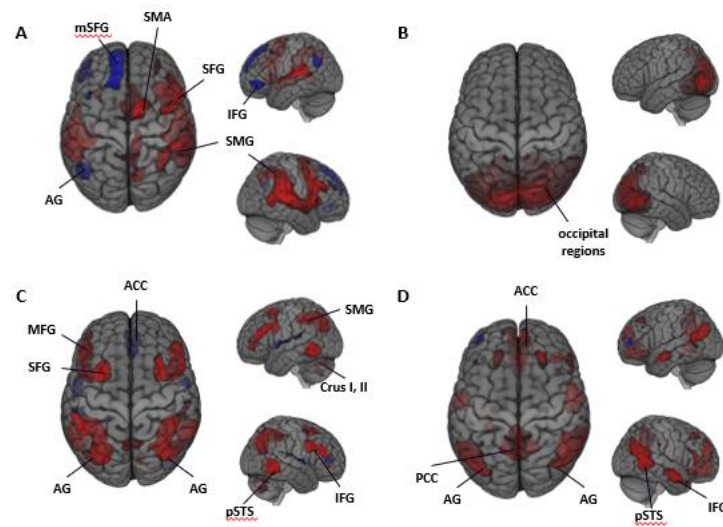


Figure 2.21 [Extended Data 1] Other brain states identified through CAP analysis

Brain states are shown after spatial standardization. **(A)** State 2 is shown. Major clusters are seen extending from the supramarginal gyrus (SMG) to the inferior, middle and superior frontal gyri. This is more prominent on the right. Activations are also observed on the supplemental motor area (SMA). Deactivations are seen on the medial superior frontal gyrus (mSFG) and angular gyrus (AG). **(B)** Visual state, with a large cluster encompassing superior, middle, and inferior occipital gyrus. **(C)** Posterior state, marked by activations spreading on the inferior parietal lobule with focus on the AG. Activations are also seen on the middle and superior frontal gyrus, and on the cerebellum Crus I and II. The anterior cingulate cortex is seen deactivated in this state **(D)** The DMN+ state is marked by activations on medial superior frontal gyrus, and anterior cingulate cortex, as well on the posterior middle cingulate cortex and angular gyri.

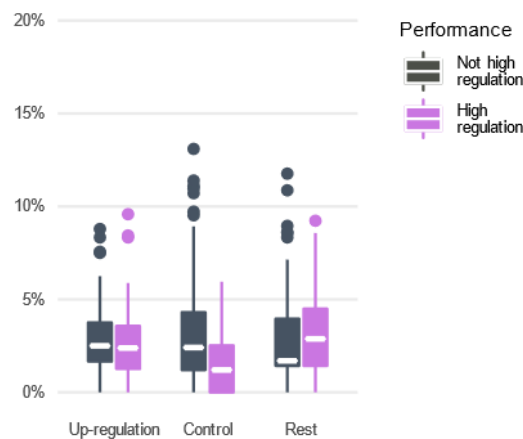


Figure 2.22 [Extended Data 2] Occurrences for the posterior state during NF runs

Linear mixed model analysis of states' occurrences identified a triple interaction between condition, high-regulation performance, and state (p -value < 0.001). Subsequent analysis revealed that the occurrences of the posterior state were significantly lower in the control condition of high-regulation performance runs, as compared to the same condition in runs that did not achieve a high-regulation performance (p -value < 0.05).

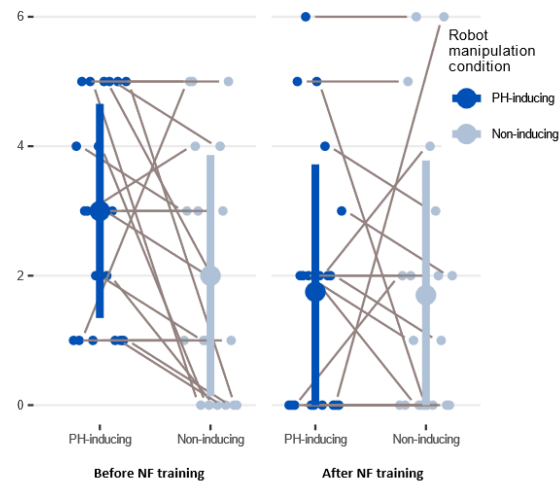


Figure 2.24 [Extended Data 3] PE ratings before and after NF training

Raw ratings for the PE induction question are shown for both asynchronous and synchronous robot manipulation condition, before neurofeedback training (left) and after neurofeedback training (right). Cumulative linear mixed model analysis showed no statistical difference between the conditions in both before and after neurofeedback training.

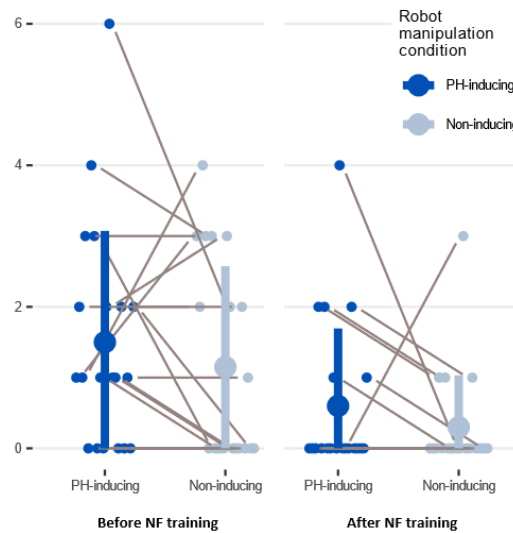


Figure 2.23 [Extended Data 4] Loss of Agency ratings before and after NF training

Raw ratings for the Loss of Agency question are shown for both asynchronous and synchronous robot manipulation condition, before neurofeedback training (left) and after neurofeedback training (right). Cumulative linear mixed model analysis showed no statistical difference between the conditions in both before and after neurofeedback training.

2.2.8 Supplemental Information

2.2.8.1 CAP Analysis

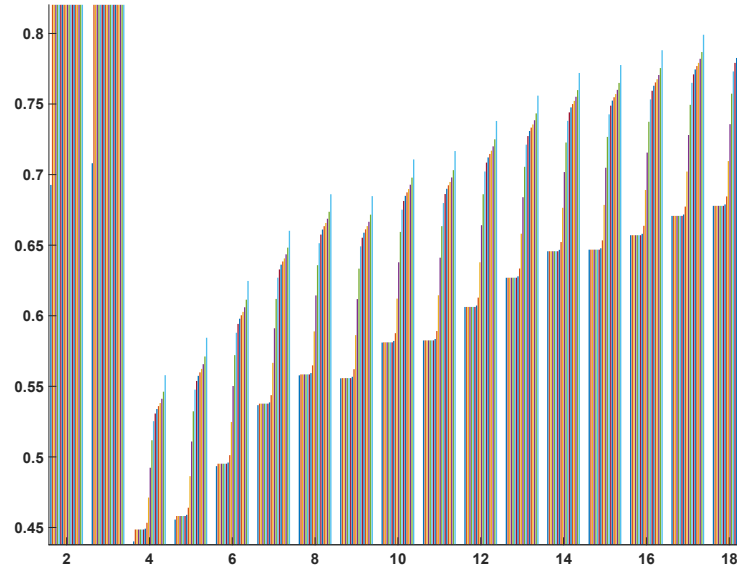


Figure 2.26 [S1] Stability of consensus clustering for different number of centroids (k)

A stability measure computed as the inverse of the number of ambiguously clustered pairs (PAC), is shown for different number of centroids. Different colours represent different consensus thresholds, used when computing the PAC. The number of centroids was selected at 8, given that besides the general tendency of Consensus Clustering to improve with increasing k (Monti et al., 2003), after $k = 8$ we observe a slight decrease in stability, indicating that $k = 8$ had in fact an increase in stability outside of the general incrementing tendency.

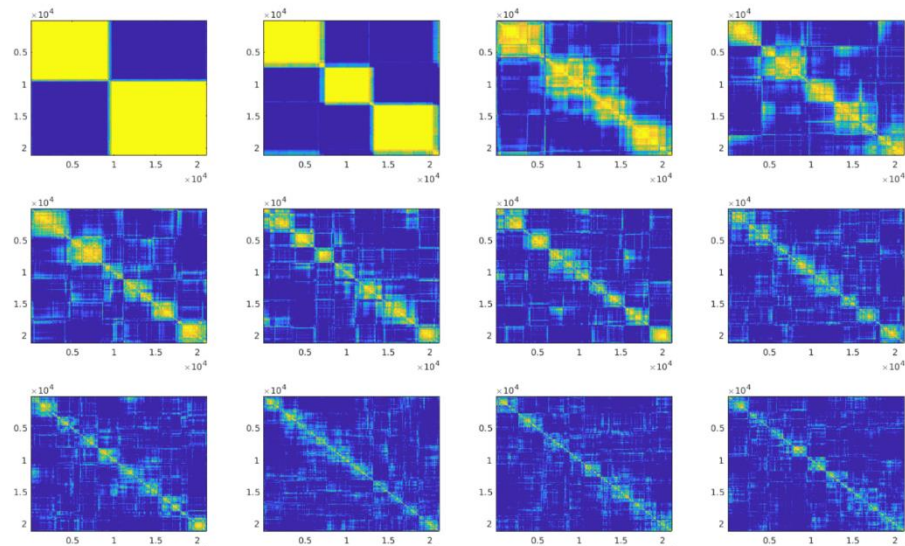


Figure 2.25 [S2] Consensus clustering for different number of centroids

A stability measure computed as the inverse of the number of ambiguously clustered pairs (PAC), is shown for different number of centroids. Different colours represent different consensus thresholds, used when computing the PAC. The number of centroids was selected at 8, given that besides the general tendency of Consensus Clustering to improve with increasing k (Monti et al., 2003), after $k = 8$ we observe a slight decrease in stability, indicating that $k = 8$ had in fact an increase in stability outside of the general incrementing tendency.

Part I: Identifying and modulating neural correlates of PH

Table 2.9 [S1] Cluster peaks for the extracted CAPs

Brain state	Main Regions (AAL) [$Z > 1.75$ & > 100 Voxels (<i>Cerebellum voxels</i> > 50)]	# Voxels	BA	Peak MNI coordinates		
				x	y	z
PH	[Deactivation] Bilateral: middle occipital gyrus; lingual; cuneus; precuneus; inferior occipital gyrus	8056	17 / 18 / 19	24	-86	26
	[Activation] Right: inferior frontal opercularis, inferior frontal triangularis, precentral gyrus, middle frontal gyrus, and insula	2373	9 / 44 / 45 / 46	50	18	26
	[Activation] Left: inferior frontal opercularis, inferior frontal triangularis, precentral	893	9 / 44	-50	10	28
	[Activation] Right middle temporal gyrus (posterior superior temporal sulcus)	131	21 / 22	58	-50	2
	[Activation] Right: Supramarginal gyrus, inferior parietal lobule	469	40	61	-31	26
	[Activation] Left: supramarginal gyrus, inferior parietal lobule, superior temporal gyrus	482	40	-58	-36	27
	[Activation] Bilateral: Supplemental Motor Area, frontal superior medial gyrus *	385	6 / 8	4	12	54
2 ⁺	[Activation] Right: superior temporal gyrus, supramarginal gyrus, inferior frontal opercularis, middle temporal gyrus, inferior frontal opercularis, inferior parietal, middle frontal gyrus, rolandic opercularis, superior temporal pole, angular gyrus	7367	9 / 13 / 21 / 22 / 40 / 42 / 45	46	18	28
	[Deactivation] Left inferior frontal orbital gyrus	324	11 / 47	-48	40	-8
	[Activation] Left: superior temporal gyrus, supramarginal gyrus, rolandic operculum, middle temporal gyrus	1887	13 / 22 / 40 / 41 / 42	-58	-36	20
	[Deactivation] Left: superior frontal gyrus, superior medial gyrus	738	8 / 9 / 10	-12	42	50
	[Deactivation] Left: angular gyrus	419	39	-46	-66	30
	[Activation] Right: precuneus, middle cingulate gyrus	708	7	10	-68	44
	[Activation] Bilateral supplemental motor area, middle cingulate gyrus	1061	6 / 32	4	10	50
TASK	[Activation] Left Cerebellum Crus I, II	80	-	-10	-80	-30
	[Activation] Right: middle frontal gyrus, inferior frontal triangularis, superior frontal gyrus, superior frontal medial gyrus, inferior frontal opercularis, inferior frontal orbital gyrus, anterior cingulate cortex, insula, posterior OFC	7427	6 / 8 / 9 / 10 / 11 / 45 / 46 / 47	50	20	26
	[Activation] Left: Insula, inferior frontal orbital gyrus	246	47	-30	22	-6

Part I: Identifying and modulating neural correlates of PH

	[Activation] Right: middle temporal gyrus, angular gyrus, inferior parietal gyrus, superior temporal gyrus, supramarginal gyrus	3281	21 / 22 / 40	62	-28	-6
	[Activation] Left: middle temporal gyrus	167	21 / 22	-62	-26	-6
	[Activation] Left: superior medial frontal gyrus	366	6 / 8 / 9	-2	42	34
	[Deactivation]: Left: Precuneus, superior parietal gyrus, postcentral gyrus, supramarginal gyrus, inferior parietal gyrus	1754	1 / 2 / 3 / 4 / 5 / 6 / 7 / 40	-56	-26	44
	[Activation] Bilateral middle cingulate cortex	191	31	4	-36	36
	[Deactivation] Right: postcentral *	279	3	56	-20	44
	[Deactivation] Bilateral supplemental motor area *	160	6 / 24	2	-10	54
	[Deactivation] Right: precuneus, superior parietal gyrus, postcentral gyrus	453	7	12	-58	60
TASK+	[Deactivation] Right: middle temporal gyrus	403	21	64	-26	-8
	[Deactivation] Right: middle frontal gyrus, inferior frontal orbital	471	10 / 47	48	42	-12
	[Deactivation] Left: middle temporal gyrus	345	21	-62	-32	-8
	[Activation] Right: middle temporal gyrus, inferior temporal gyrus	827	37 / 39	56	-54	0
	[Activation] Left: middle temporal gyrus, inferior temporal gyrus	310	37	-56	-62	-2
	[Deactivation] Bilateral: superior medial frontal gyrus *	1743	8 / 9 / 32	4	42	36
	[Activation] Right: supramarginal gyrus, postcentral gyrus, rolandic operculum	1583	2 / 40	60	-28	38
	[Activation] Left: supramarginal gyrus, inferior parietal gyrus, postcentral gyrus	882	2 / 40	-62	-32	32
	[Deactivation] Left: angular gyrus, inferior parietal gyrus	828	40	-46	-58	46
	[Deactivation] Right: angular gyrus, inferior parietal gyrus	723	40	52	-52	42
	[Deactivation] Bilateral middle cingulate cortex	170	31	-8	-44	36
	[Deactivation] Right Middle frontal gyrus	322	8 / 9	40	16	44
	[Deactivation] Left middle frontal gyrus	213	8 / 9	-44	18	42
	[Activation] Bilateral: Precuneus, superior parietal gyrus *	704	7	-10	-56	58
	[Activation] Right: superior frontal gyrus, supplemental motor area	588	6	26	-6	62
	[Activation] Left: superior frontal gyrus	177	6	-26	-6	56
Visual	[Activation] Bilateral: Middle occipital, lingual gyrus, calcarine, middle temporal gyrus, cuneus, fusiform gyrus,	15006	7 / 17 / 18 / 19 / 30 /	56	-56	0

Part I: Identifying and modulating neural correlates of PH

	superior occipital, inferior temporal gyrus, inferior occipital, superior parietal		31 / 37 / 39			
DMN-	[Activation] Right: inferior frontal triangularis, inferior frontal opercularis, precentral gyrus, middle frontal gyrus	2662	6 / 9 / 45 / 46	50	18	26
	[Activation] Left: inferior frontal triangularis, inferior frontal opercularis	564	9	-48	12	28
	[Deactivation] Left: anterior cingulate cortex. Bilateral: superior medial frontal	485	10 / 32	-4	54	2
	[Activation] Right: superior temporal gyrus, middle temporal gyrus	565	22	50	-24	-2
	[Deactivation] Bilateral: precuneus, middle cingulate cortex, posterior cingulate cortex, cuneus, calcarine	3902	7 / 23 / 31	0	-62	38
	[Activation] Left: Superior temporal gyrus	203	22 / 42	-64	-36	10
	[Deactivation] Right: Angular gyrus, middle occipital gyrus	1038	39	48	-64	32
	[Deactivation]: Left: angular gyrus, middle occipital gyrus	1294	39	-40	-72	36
	[Deactivation] Left: superior frontal gyrus	417	8	-26	26	52
	[Deactivation] Right: superior frontal gyrus	295	8	24	38	50
	[Activation] Left precentral gyrus	170	4 / 6	-50	-8	52
	[Activation] Bilateral supplemental motor area *	315	6	4	8	62
Post.	[Activation] Right: Cerebellum crus I, II; Cerebellum 6, 7b	415	-	30	-64	-30
	[Activation] Right: Cerebellum crus I, II	107	-	10	-78	-28
	[Activation] Right: Middle temporal gyrus, inferior temporal gyrus	950	21 / 37	58	-52	-2
	[Activation] Left: Middle temporal gyrus, inferior temporal gyrus	871	21 / 37	-56	-58	-8
	[Activation] Left: inferior frontal triangularis, middle frontal gyrus, inferior frontal opercularis	1116	9 / 46	-45	24	23
	[Deactivation] Left superior temporal gyrus	333	22 / 41	-54	4	0
	[Deactivation] Right rolandic opercularis	187	22	58	0	4
	[Deactivation] Bilateral anterior cingulate cortex	432	32	4	44	6
	[Activation] Right: inferior frontal triangularis, inferior frontal opercularis, middle frontal gyrus	1146	9 / 46	48	34	20
	[Activation] Right: Angular gyrus, inferior parietal gyrus, middle occipital gyrus, superior parietal gyrus, supramarginal gyrus	2242	7 / 39 / 40	34	-70	40
	[Activation] Left: Angular gyrus, inferior parietal gyrus, middle occipital gyrus, superior parietal gyrus, precuneus	2204	7 / 19 / 40	-30	-74	40

Part I: Identifying and modulating neural correlates of PH

	[Activation] Left: Middle frontal gyrus, superior frontal gyrus	432	6 / 8	-26	12	62
	[Activation] Right: Middle frontal gyrus, superior frontal gyrus	357	6 / 8	30	10	60
DMN [†]	[Activation] Right: Middle temporal gyrus, superior temporal gyrus	778	21	58	-4	-16
	[Activation] Left: Middle temporal gyrus, superior temporal gyrus	473	21	-58	-8	-14
	[Activation] Bilateral: Medial frontal orbital gyrus, medial superior frontal gyrus, anterior cingulate cortex	2198	9 / 10 / 11 / 32	2	52	-10
	[Activation] Right: posterior OFC	134	11 / 47	38	34	-14
	[Activation] Right: middle temporal gyrus, angular gyrus, superior temporal gyrus, middle occipital gyrus, inferior temporal gyrus	2549	22 / 39	58	-52	0
	[Activation] Right: inferior frontal triangularis	179	45 / 46	52	36	8
	[Activation] Left: Middle temporal gyrus, middle occipital gyrus, angular gyrus	1434	19 / 39	-42	-74	30
	[Deactivation] Left: Middle frontal gyrus	120	10	-40	56	8
	[Activation] Bilateral: precuneus, middle cingulate cortex, posterior middle cingulate cortex, calcarine	2568	7 / 23 / 30 / 31	2	-60	26
	[Activation] Left superior frontal gyrus	144	8 / 9	-20	34	46
	[Activation] Right superior frontal gyrus	188	8	24	36	48

Main regions were considered regions when exceeding a z-score of 1.75, and had more than 100 voxels if in cortical regions, or 50 voxels when in the cerebellum.

* Two clusters are contained in these medial clusters which were not communicating.

[†] We consider that this state mostly captured noise.

Part I: Identifying and modulating neural correlates of PH

2.2.8.2 Analysis of transition probabilities during NF runs

Table 2.10 [2] Results of the transition probability analysis for NF runs

Parameter	χ^2	DoF	p-value
Intercept	935.29	1	< 0.001
Initial State	84.83	8	< 0.001
Final State	801.46	8	< 0.001
Condition	0.25	2	0.883
HRP	10.08	1	0.002
Initial State : Final State	263.29	64	< 0.001
Initial State : Condition	19.23	16	0.257
Final State : Condition	1.46	16	0.999
Initial State: HRP	9.01	8	0.341
Final State: HRP	8.35	8	0.400
Condition: HRP	0.14	2	0.931
Initial State : Final State : Condition	55.63	128	0.999
InitialState : FinalState: HRP	23.48	64	0.999
InitialState : Condition: HRP	40.89	16	< 0.001
FinalState : Condition: HRP	0.37	16	0.999
InitialState : FinalState : Condition: HRP	89.28	128	0.996

This model included parameters for initial state, final state, condition, and HRP. Highlighted in green is the triple interaction which motivates the post-hoc decomposition into more specific models.

Table 2.11 [S3] Intermediate results for the transition probability analysis of NF runs

Condition	p-value intercept	p-value final	p-value initial	p-value HRP	p-value final: initial	p-value final:HRP	p-value initial:HRP	p-value final: initial: HRP
Upregulation	< 0.001	< 0.001	< 0.001	< 0.001	< 0.001	0.091	0.028	0.985
Control	< 0.001	< 0.001	< 0.001	0.004	< 0.001	0.664	0.493	0.985
Rest	< 0.001	< 0.001	< 0.001	0.008	< 0.001	0.679	< 0.001	0.985

This decomposition for the levels of condition was performed following the previous three-way interaction, and is used to reduce the number of multiple comparisons of the final models. Highlighted in green are the interactions that motivate the continued decomposition of the models (and supported by the previous three-way interaction). Note that while the control condition included an interaction between initial and final state, this interaction was not supported by the higher level interaction, for continued decomposition for this condition.

Part I: Identifying and modulating neural correlates of PH

Table 2.12 [S4] Final analysis for the transition probabilities of different brain states during up-regulation or rest for NF runs

Condition	Initial State	p-value intercept	p-value final	p-value HRP	p-value final:HRP
Up-regulation	state ⁰	<0.001	<0.001	<0.001	0.014
Up-regulation	PH	<0.001	<0.001	<0.001	0.007
Up-regulation	2	<0.001	<0.001	0.012	0.412
Up-regulation	TASK ⁻	<0.001	<0.001	<0.001	<0.001
Up-regulation	TASK ⁺	<0.001	<0.001	0.001	0.412
Up-regulation	Visual	<0.001	<0.001	0.498	0.966
Up-regulation	DMN ⁻	<0.001	<0.001	<0.001	<0.001
Up-regulation	Post.	<0.001	<0.001	0.003	0.225
Up-regulation	DMN ⁺	<0.001	<0.001	<0.001	0.019
Rest	state ⁰	<0.001	<0.001	<0.001	0.060
Rest	PH	<0.001	<0.001	<0.001	0.060
Rest	2	<0.001	<0.001	0.713	0.922
Rest	TASK ⁻	<0.001	<0.001	<0.001	0.053
Rest	TASK ⁺	<0.001	<0.001	0.008	0.751
Rest	Visual	<0.001	<0.001	0.082	0.92
Rest	DMN ⁻	<0.001	<0.001	<0.001	0.001
Rest	Post.	<0.001	<0.001	0.078	0.922
Rest	DMN ⁺	<0.001	<0.001	0.386	0.556

Highlighted in green are results for states and conditions, in which both an effect of HRP was observed, alongside with a significant effect of the interaction between HRP and final state.

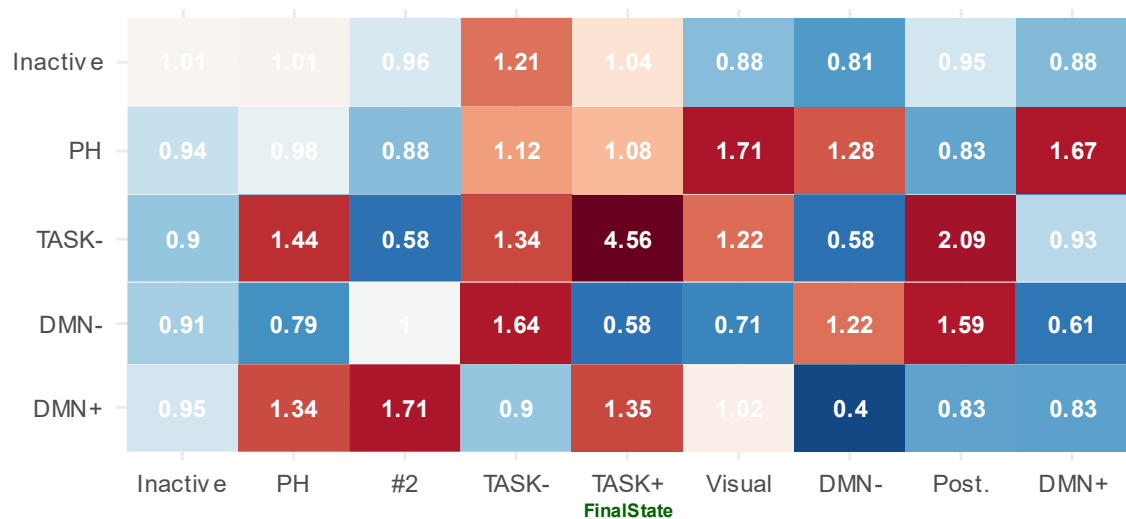


Figure 2.27 [S3] Change in transition probabilities when comparing HRP versus non-HRP runs during up-regulation. Increase in transition probabilities, when comparing HRP runs and non HRP-runs in the up-regulation condition is shown here for the states that observed both an HRP effect and an interaction between HRP and final state (inactive, PH, TASK, DMN-, DMN+).

Part I: Identifying and modulating neural correlates of PH

2.2.8.3 Analysis of transition probabilities during transfer runs

Table 2.13[S5] Results for the unrestricted transition probability analysis during transfer runs

Parameter	χ^2	DoF	p-value
Intercept	464.50	1	<0.001
Initial State	49.09	8	<0.001
Final State	393.23	8	<0.001
Condition	0.31	2	0.857
HRP	29.55	1	<0.001
Initial : Final	107.49	64	<0.001
Initial State : Condition	127.56	16	<0.001
Final State : Condition	0.70	16	0.999
Initial State: HRP	8.08	8	0.426
Final State : HRP	25.03	8	0.002
Condition : HRP	0.19	2	0.911
Initial State : Final State : Condition	184.76	128	<0.001
Initial State : Final State : HRP	41.38	64	0.988
Initial State : Condition : HRP	36.87	16	0.002
Final State: Condition: HRP	0.43	16	0.999
Initial State : Final State : Condition : HRP	88.15	128	0.997

This model included parameters for initial state, final state, condition, and HRP. Highlighted in green is the triple interaction which supports the post-hoc decomposition into more specific models.

Table 2.14 [S6] Restricted analysis of transition probabilities during transfer runs

Condition	Initial State	p-value intercept	p-value final	p-value HRP	p-value HRP: final
Upregulation	state ⁰	<0.001	<0.001	<0.001	<0.001
Upregulation	PH	<0.001	<0.001	<0.001	<0.001
Upregulation	TASK ⁻	<0.001	<0.001	0.014	0.553
Upregulation	DMN ⁻	<0.001	<0.001	0.001	0.087
Upregulation	DMN ⁺	<0.001	<0.001	0.001	0.087

This analysis was done specifically for the up-regulation condition, and for the states where significant changes were observed during NF runs. Two states, the state⁰ and PH, show a significant difference in transition probabilities due to HRP and also an interaction between HRP and final state.

Part I: Identifying and modulating neural correlates of PH

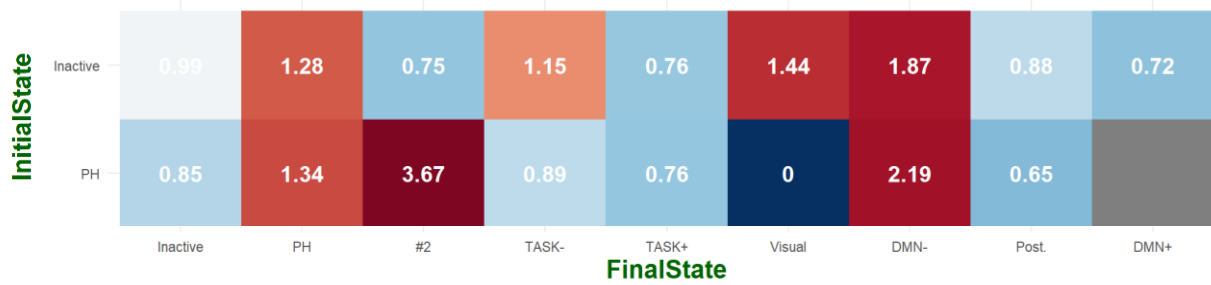


Figure 2.28 [S4] Change in transition probabilities when comparing the up-regulation condition between HRP and non-HRP transfer runs

Only the inactive state and the PH state showed a significant difference in transition probabilities between the HRP and non-HRP transfer runs. Here the difference between HRP and non-HRP runs is shown. This pattern did not fully mimic the one observed in NF runs, albeit an increase from state⁰ to TASK⁻ and from the PH state to DMN⁻. We do note that the number of transitions available for transfer runs is much less than for NF runs (9 times lower), and hence these transition matrixes are less reliable.

2.2.8.4 Analysis of occurrences during NF runs

Table 2.15 [S7] Results of the analysis of occurrences of brain states during NF runs

Parameter	χ^2	DoF	p-value
Intercept	64467.58	1	< 0.001
Condition	20.55	2	< 0.001
State	52140.94	8	< 0.001
HRP	9	1	0.592
Condition : State	141.81	16	< 0.001
Condition : HRP	13.27	2	0.001
State : HRP	4.33	8	0.826
Condition : State : HRP	49.11	16	< 0.001

This model included parameters for initial state, final state, condition, and HRP. Highlighted in green is the triple interaction which motivates the post-hoc decomposition into more specific models.

Table 2.16 [S8] Intermediate results for the analysis of occurrences of brain states during NF runs

Condition	p-value intercept	p-value state	p-value HRP	p-value state : HRP
Up-regulation	< 0.001	< 0.001	0.411	0.252
Control	< 0.001	< 0.001	< 0.001	< 0.001
Rest	< 0.001	< 0.001	0.048	0.252

This decomposition for the levels of condition was performed following the previous three-way interaction, and is used to reduce the number of multiple comparisons of the final models. Highlighted in green is the interaction that motivate the continued decomposition of the models (and supported by the previous three-way interaction).

Part I: Identifying and modulating neural correlates of PH

Table 2.17 [S9] Final results for the analysis of occurrences of brain states during NF runs

Condition	State	p-value HRP
Control	state ⁰	0.190
Control	PH	0.026
Control	2	0.367
Control	TASK ⁻	0.190
Control	TASK ⁺	0.048
Control	Visual	0.190
Control	DMN ⁻	0.366
Control	Post.	0.018
Control	DMN ⁺	0.366

Highlighted in green are the states which changed their occurrences based on HRP during the control condition.

2.2.8.5 Analysis of occurrences during transfer runs

To analyze if there was sustainment of the neural regulation mechanisms developed during the NF training runs to the transfer runs, we performed first an unrestricted LMM analysis with parameters condition, state and HRP. If a meaningful interaction was found, we then proceeded with a post-hoc analysis restricted to the neural strategy found in NF runs.

Table 2.18 [S10] Results for the unrestricted analysis of occurrences of brain states during transfer runs

Parameter	χ^2	DoF	p-value
Intercept	9756.10	1	<0.001
Condition	16.02	2	<0.001
State	8031.42	8	<0.001
HRP	0.27	1	0.61
Condition : State	28.38	16	0.028
Condition : HRP	7.32	2	0.027
State : HRP	0.92	8	0.999
Condition : State : HRP	11.50	16	0.778

This model included parameters for state, condition, and HRP. Highlighted in green are the triple interaction which motivated the post-hoc decomposition into more specific models.

Table 2.19 [S11] Results for the restricted analysis of the occurrences of brain states during transfer runs

Condition	State	p-value HRP
Control	PH	0.301
Control	TASK ⁺	0.600
Control	Post.	0.423

This analysis focused on the states of PH, TASK, and Posterior, which had significantly different occurrences in the control condition of NF runs, when comparing HRP and non-HRP runs.

Part I: Identifying and modulating neural correlates of PH

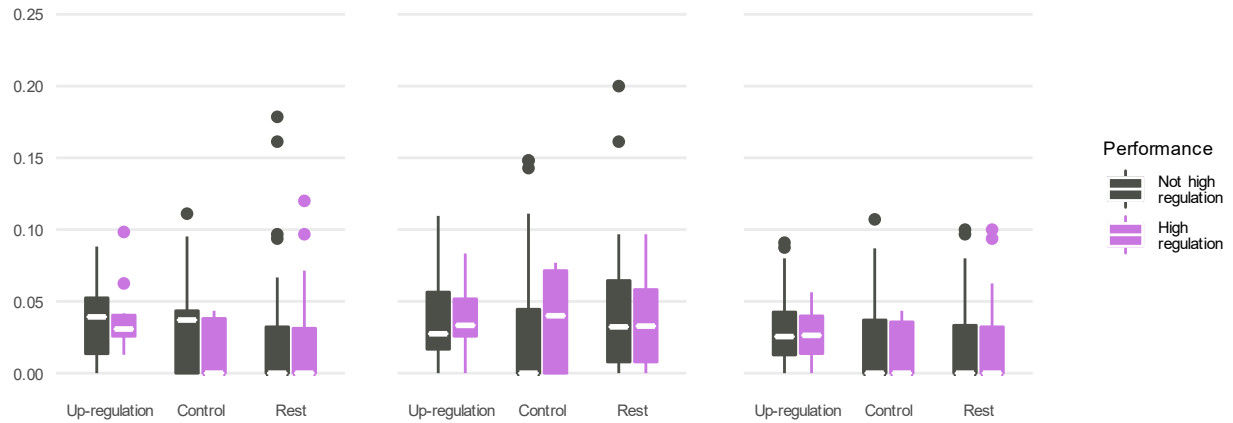


Figure 2.29 [S5] Occurrences for the different states during transfer runs (HRP vs non-HRP runs)
(A) PH-state. (B) Task⁺ state (C) Posterior state.

2.2.8.6 Behavioral differences due to training in control questions

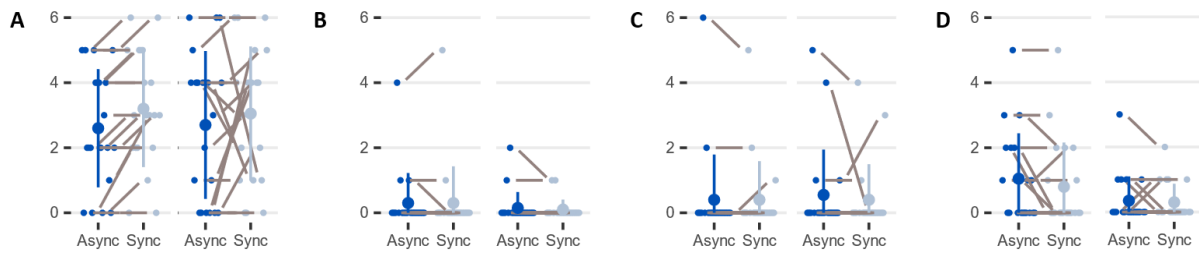


Figure 2.30 [S6] Ratings for the control questions of the PH-questionnaire

Raw ratings for different control questions are shown. **(A)** Perceived self-touch **(B)** Induction of PH in front **(C)** Sensation of having two bodies **(D)** Generalised anxiety. None of these sensations showed any significant change due to synchrony after or before training. For the induction of PH at the front and the two bodies question, the model could not converge due to the large amount of zeros.

Part I: Identifying and modulating neural correlates of PH

2.2.8.7 Analysis of differential occurrences between day 5 and day 1

Table 2.20 [S12] Unrestricted analysis of differential occurrences of PH-task sessions

Parameter	χ^2	DoF	p-value
Intercept	0.037	1	0.850
Condition	0.457	1	0.499
State	3.064	8	0.930
#HRP	2.771	1	0.096
PH	0.653	1	0.419
Condition : State	3.295	8	0.915
Condition : #HRP	6.598	1	0.010
State : #HRP	5.221	8	0.734
Condition : PH	2.730	1	0.098
State : PH	7.570	8	0.476
#HRP : PH	2.825	1	0.092
Condition : State : #HRP	10.160	8	0.254
Condition : State : PH	10.205	8	0.251
Condition : #HRP : PH	17.417	1	< 0.001
State : #HRP : PH	10.703	8	0.219
Condition : State : #HRP : PH	25.568	8	0.001

Linear mixed model analysis of the various parameters integrating the higher-level differential occurrences model for the PH-task runs, according to the Wal-chi square test.

Table 2.21 [S13] Intermediate model for differential occurrences of PH-task sessions

PH	p-value intercept	p-value condition	p-value #HRP	p-value state	p-value condition: #HRP	p-value condition: state	p-value #HRP: state	p-value condition: #HRP : state
Not sensitive	0.863	0.781	0.256	0.959	0.017	0.992	0.822	0.462
Sensitive	0.440	0.018	0.357	0.417	0.006	0.010	0.374	0.045

Intermediate model comparing the effects of the parameters: state, number of HRP runs obtained during NF training (#HRP) and condition, for different levels of sensitivity to PH induction on the last day of NF training. This intermediate model was used to reduce the number of multiple comparisons following the quadruple interaction detected above.

Part I: Identifying and modulating neural correlates of PH

Table 2.22 [S14] Intermediate model 2 for differential occurrence of PH-induction sessions

PH	State	p-value intercept	p-value condition	p-value #HRP	p-value condition:#HRP
Sensitive	state ⁰	0.672	0.273	0.696	0.179
Sensitive	PH	0.274	0.033	0.201	0.035
Sensitive	2	0.462	0.545	0.696	0.403
Sensitive	TASK ⁻	0.949	0.694	0.696	0.746
Sensitive	TASK ⁺	0.231	0.185	0.201	0.068
Sensitive	Visual	0.232	0.224	0.201	0.169
Sensitive	DMN ⁻	0.232	0.312	0.201	0.403
Sensitive	Post.	0.462	0.272	0.429	0.537
Sensitive	DMN ⁺	0.949	0.545	0.539	0.413

Intermediate model comparing the effects of the parameters: number of HRP runs obtained during NF training (#HRP) and condition, for different states, in participants sensitive to PH at the last day. This intermediate model was used to reduce the number of multiple comparisons following the quadruple interaction detected above.

Table 2.23 [S15] Final model assessing differential occurrences of PH-induction sessions

PH	State	Condition	p-value #HRP
Sensitive	PH	PH-inducing	0.166
Sensitive	PH	non-inducing	0.166

Final model assessing the effect of the number of HRP runs obtained during NF training on the occurrences of the PH-network, in participants sensitive to PH and in the two conditions of the PH-task.

Table 2.24 [S16] Control analysis for differential occurrences of passivity experiences (PE) – unrestricted model

Parameter	χ^2	DoF	p-value
Intercept	3.07	1	0.080
Condition	7.56	1	0.006
State	4.92	8	0.766
#HRP	0.80	1	0.370
PE	4.17	1	0.041
Condition : State	10.62	8	0.224
Condition : #HRP	1.69	1	0.194
State : #HRP	2.95	8	0.937
Condition : PE	4.26	1	0.039
State : PE	7.421	8	0.492
#HRP : PE	0.05	1	0.817
Condition : State : #HRP	5.07	8	0.750
Condition : State : PE	7.79	8	0.453
Condition : #HRP : PE	1.26	1	0.262
State : #HRP : PE	5.03	8	0.754
Condition : State : #HRP : PE	4.29	8	0.830

Part I: Identifying and modulating neural correlates of PH

Table 2.25 [S17] Control analysis for differential occurrences of PE - restricted model

PE	p-value intercept	p-value condition	p-value #HRP	p-value state	p-value condition: #HRP	p-value condition: state	p-value #HRP: state	p-value condition: #HRP : state
Not sensitive	0.122	0.046	0.415	0.851	0.302	0.752	0.965	0.907
Sensitive	0.122	0.046	0.415	0.533	0.302	0.056	0.583	0.355

Table 2.26 [S18] Control analysis for differential occurrences using Loss of Agency (LoA)

Parameter	χ^2	DoF	p-value
Intercept	0.28	1	0.595
Condition	2.54	1	0.111
State	3.71	8	0.882
#HRP	3.56	1	0.059
LoA	0.50	1	0.480
Condition : State	8.32	8	0.402
Condition : #HRP	4.35	1	0.037
State : #HRP	5.36	8	0.718
Condition : LoA	0.41	1	0.523
State : LoA	5.24	8	0.731
#HRP : LoA	4.90	1	0.027
Condition : State : #HRP	11.09	8	0.197
Condition : State : LoA	4.17	8	0.841
Condition : #HRP : LoA	11.32	1	0.001
State : #HRP : LoA	10.95	8	0.204
Condition : State : #HRP : LoA	19.11	8	0.014

Table 2.27 [S19] Control analysis (using LoA) for differential occurrences of PH-induction sessions - restricted model 1

LoA	p-value intercept	p-value condition	p-value #HRP	p-value state	p-value condition: #HRP	p-value condition: state	p-value #HRP: state	p-value condition: #HRP : state
Not sensitive	0.626	0.328	0.111	0.926	0.058	0.712	0.808	0.530
Sensitive	0.497	0.066	0.111	0.810	0.023	0.712	0.315	0.530

Part I: Identifying and modulating neural correlates of PH

Table 2.28 [S20] Control analysis (using LoA) for differential occurrences of PH-induction sessions - restricted model 2

LoA	Condition	p-value intercept	p-value #HRP	p-value state	p-value #HRP: state
Sensitive	PH-inducing	0.278	0.134	0.921	0.233
Sensitive	Non-inducing	0.278	0.013	0.921	0.178

2.2.8.1 Analysis of differential transition probabilities between day 5 and day 1

Parameter	χ^2	DoF	p-value
Intercept	0.006	1	0.938
Initial	6.26	8	0.618
Final	0.12	8	0.999
Condition	0.01	1	0.920
#HRP	0	1	0.988
PH	0.007	1	0.932
Initial : Final	54.37	64	0.799
Initial : Condition	18.01	8	0.021
Final : Condition	0.11	8	0.999
Initial : #HRP	9.02	8	0.341
Final : #HRP	0.14	8	0.999
Condition : #HRP	0.09	1	0.765
Initial : PH	10.02	8	0.264
Final : PH	0.14	8	0.999
Condition : PH	0.059	1	0.808
#HRP : PH	0	1	0.977
Initial : Final : Condition	57.72	64	0.697
Initial : Final : #HRP	65.77	64	0.415
Initial : Condition : #HRP	9.39	8	0.310
Final : Condition : #HRP	0.16	8	0.999
Initial : Final : PH	68.54	64	0.326
Initial : Condition : PH	9.52	8	0.301
Final : Condition : PH	0.19	8	0.999
Initial : #HRP : PH	11.34	8	0.183
Final : #HRP : PH	0.15	8	0.999
Condition : #HRP : PH	0.15	1	0.701
Initial : Final : Condition : #HRP	57.17	64	0.715
Initial : Final : Condition : PH	47.29	64	0.942
Initial : Final : #HRP : PH	76.01	64	0.145
Initial : Condition : #HRP : PH	14.60	8	0.067
Final : Condition : #HRP : PH	0.28	8	0.999
Initial : Final : Condition : #HRP : PH	58.05	64	0.686

Table 2.29 [S21] Analysis of differential transition probabilities of PH-induction days - unrestricted

Part I: Identifying and modulating neural correlates of PH

Table 2.30 [S22] Final model of differential transition probabilities of PH-induction days

Condition	Initial State	p-value intercept	p-value final	p-value #HRP	p-value PH	p-value final:#HRP	p-value final:PH	p-value #HRP:PH	p-value #HRP:final:PH
Async	state ⁰	0.862	0.995	0.941	0.924	0.779	0.839	0.869	0.701
Async	PH	0.636	0.995	0.198	0.924	0.926	0.999	0.439	0.942
Async	2	0.928	0.995	0.652	0.062	0.993	0.635	0.544	0.715
Async	TASK ⁻	0.911	0.995	0.330	0.924	0.680	0.839	0.439	0.647
Async	TASK ⁺	0.583	0.995	0.655	0.924	0.993	0.511	0.880	0.062
Async	Visual	0.928	0.995	0.652	0.731	0.993	0.999	0.205	0.834
Async	DMN ⁻	0.911	0.995	0.710	0.559	0.993	0.999	0.439	0.715
Async	Post.	0.583	0.995	0.198	0.924	0.680	0.999	0.439	0.645
Async	DMN ⁺	0.636	0.409	0.652	0.924	0.512	0.635	0.529	0.645
Sync	state ⁰	0.583	0.409	0.020	0.327	0.0799	0.635	0.002	0.062
Sync	PH	0.762	0.995	0.330	0.924	0.926	0.839	0.335	0.645
Sync	2	0.583	0.995	0.452	0.924	0.993	0.9998	0.439	0.942
Sync	TASK ⁻	0.862	0.995	0.237	0.924	0.993	0.9998	0.205	0.715
Sync	TASK ⁺	0.583	0.995	0.319	0.001	0.080	0.233	0.048	0.645
Sync	Visual	0.583	0.995	0.197	0.924	0.993	0.999	0.439	0.942
Sync	DMN ⁻	0.001	0.060	0.198	0.970	0.779	0.999	0.335	0.715
Sync	Post.	0.991	0.995	0.198	0.731	0.6780	0.999	0.869	0.942
Sync	DMN ⁺	0.862	0.995	0.941	0.731	0.993	0.999	0.205	0.715

Part I: Identifying and modulating neural correlates of PH

Table 2.31 [S23] Control analysis (PE) for differential transition probabilities - unrestricted

Parameter	χ^2	DoF	p-value
Intercept	0.084	1	0.771
Initial	9.813	8	0.278
Final	0.16	8	1
Condition	0.029	1	0.865
#HRP	0.023	1	0.879
PE	0.078	1	0.779
Initial : Final	66.693	64	0.385
Initial : Condition	33.68	8	0
Final : Condition	0.094	8	1
Initial : #HRP	13.091	8	0.109
Final : #HRP	0.144	8	1
Condition : #HRP	0.014	1	0.905
Initial : PE	22.821	8	0.004
Final : PE	0.128	8	1
Condition : PE	0.063	1	0.802
#HRP : PE	0.058	1	0.81
Initial : Final : Condition	77.317	64	0.123
Initial : Final : #HRP	68.374	64	0.331
Initial : Condition : #HRP	15.28	8	0.054
Final : Condition : #HRP	0.066	8	1
Initial : Final : PE	60.243	64	0.61
Initial : Condition : PE	23.678	8	0.003
Final : Condition : PE	0.123	8	1
Initial : #HRP : PE	20.278	8	0.009
Final : #HRP : PE	0.155	8	1
Condition : #HRP : PE	0	1	0.999
Initial : Final : Condition : #HRP	62.524	64	0.529
Initial : Final : Condition : PE	57.387	64	0.708
Initial : Final : #HRP : PE	66.375	64	0.395
Initial : Condition : #HRP : PE	25.052	8	0.002
Final : Condition : #HRP : PE	0.072	8	1
Initial : Final : Condition : #HRP : PE	60.92	64	0.586

Part I: Identifying and modulating neural correlates of PH

Table 2.32 [S24] Control analysis (PE) for differential transition probabilities - unrestricted

PH	Condition	Initial State	p-value intercept	p-value final	p-value #HRP	p-value final : #HRP
FALSE	Async	state ⁰	0.065	0.688	0.387	0.637
FALSE	Async	PH	0.413	0.982	0.062	0.637
FALSE	Async	2	0.609	0.982	0.739	0.98
FALSE	Async	TASK ⁻	0.465	0.688	0.114	0.382
FALSE	Async	TASK ⁺	0.301	0.982	0.617	0.98
FALSE	Async	Visual	0.738	0.982	0.625	0.987
FALSE	Async	DMN ⁻	0.738	0.982	0.78	0.98
FALSE	Async	Post.	0.411	0.457	0.086	0.382
FALSE	Async	DMN ⁺	0.411	0.505	0.739	0.596
FALSE	Sync	state ⁰	0.821	0.822	0.062	0.382
FALSE	Sync	PH	0.821	0.982	0.623	0.98
FALSE	Sync	2	0.03	0.781	0.23	0.98
FALSE	Sync	TASK ⁻	0.411	0.982	0.412	0.98
FALSE	Sync	TASK ⁺	0.003	0.173	0.034	0.063
FALSE	Sync	Visual	0.063	0.568	0.034	0.637
FALSE	Sync	DMN ⁻	0.003	0.173	0.114	0.771
FALSE	Sync	Post.	0.847	0.982	0.086	0.527
FALSE	Sync	DMN ⁺	0.712	0.982	0.904	0.98
TRUE	Async	state ⁰	0.411	0.982	0.306	0.98
TRUE	Async	PH	0.314	0.658	0.034	0.382
TRUE	Async	2	0.061	0.791	0.533	0.98
TRUE	Async	TASK ⁻	0.078	0.658	0.068	0.681
TRUE	Async	TASK ⁺	0.847	0.791	0.904	0.382
TRUE	Async	Visual	0.055	0.791	0.034	0.596
TRUE	Async	DMN ⁻	0.012	0.224	0.189	0.596
TRUE	Async	Post.	0.411	0.982	0.114	0.596
TRUE	Async	DMN ⁺	0.697	0.982	0.904	0.98
TRUE	Sync	state ⁰	0.645	0.982	0.67	0.98
TRUE	Sync	PH	0.847	0.982	0.904	0.98
TRUE	Sync	2	0.853	0.791	0.625	0.98
TRUE	Sync	TASK ⁻	0.509	0.982	0.904	0.987
TRUE	Sync	TASK ⁺	0.01	0.224	0.034	0.596
TRUE	Sync	Visual	0.903	0.982	0.739	0.98
TRUE	Sync	DMN ⁻	0.006	0.173	0.022	0.382
TRUE	Sync	Post.	0.301	0.982	0.309	0.98
TRUE	Sync	DMN ⁺	0.321	0.982	0.034	0.454

Part I: Identifying and modulating neural correlates of PH

Table 2.33 [S25] Control analysis (LoA) for differential transition probabilities - unrestricted

Parameter	χ^2	DoF	p-value
Intercept	0.11	1	0.74
Initial	16.206	8	0.04
Final	0.185	8	1
Condition	0.022	1	0.883
#HRP	0.012	1	0.913
LoA	0.098	1	0.754
Initial : Final	64.149	64	0.471
Initial : Condition	29.773	8	<0.001
Final : Condition	0.171	8	1
Initial : #HRP	15.863	8	0.044
Final : #HRP	0.064	8	1
Condition : #HRP	0.026	1	0.871
Initial : LoA	49.733	8	<0.001
Final : LoA	0.205	8	1
Condition : LoA	0.025	1	0.874
#HRP : LoA	0.019	1	0.89
Initial : Final : Condition	72.182	64	0.226
Initial : Final : #HRP	54.897	64	0.784
Initial : Condition : #HRP	15.382	8	0.052
Final : Condition : #HRP	0.159	8	1
Initial : Final : LoA	100.997	64	0.002
Initial : Condition : LoA	17.729	8	0.023
Final : Condition : LoA	0.136	8	1
Initial : #HRP : LoA	68.806	8	<0.001
Final : #HRP : LoA	0.15	8	1
Condition : #HRP : LoA	0.016	1	0.898
Initial : Final : Condition : #HRP	50.728	64	0.886
Initial : Final : Condition : LoA	53.669	64	0.818
Initial : Final : #HRP : LoA	115.202	64	<0.001
Initial : Condition : #HRP : LoA	31.455	8	<0.001
Final : Condition : #HRP : LoA	0.177	8	1
Initial : Final : Condition : #HRP : LoA	63.337	64	0.5

Part I: Identifying and modulating neural correlates of PH

Table 2.34 [S26] Control analysis (LoA) for differential transition probabilities - unrestricted

LoA	Condition	Initial State	p-value intercept	p-value final	p-value #HRP	p-value final : #HRP
FALSE	Async	state ⁰	0.022	0.384	0.525	0.999
FALSE	Async	PH	0.653	0.998	0.195	0.945
FALSE	Async	2	0.063	0.717	0.407	0.999
FALSE	Async	TASK ⁻	0.549	0.739	0.216	0.822
FALSE	Async	TASK ⁺	0.05	0.429	0.552	0.945
FALSE	Async	Visual	0.108	0.739	0.02	0.556
FALSE	Async	DMN ⁻	0.996	0.998	0.928	0.999
FALSE	Async	Post.	0.719	0.526	0.719	0.945
FALSE	Async	DMN ⁺	0.391	0.654	0.755	0.945
FALSE	Sync	state ⁰	0.639	0.739	0.071	0.3
FALSE	Sync	PH	0.895	0.896	0.407	0.945
FALSE	Sync	2	0.057	0.877	0.673	0.999
FALSE	Sync	TASK ⁻	0.92	0.998	0.13	0.945
FALSE	Sync	TASK ⁺	0.002	0.019	0.013	0.033
FALSE	Sync	Visual	0.005	0.123	0.525	0.999
FALSE	Sync	DMN ⁻	"0.005	0.237	0.116	0.945
FALSE	Sync	Post.	0.996	0.998	0.425	0.945
FALSE	Sync	DMN ⁺	0.895	0.998	0.383	0.945
TRUE	Async	state ⁰	0.384	0.717	0.572	0.552
TRUE	Async	PH	0.01	0.237	0	0.004
TRUE	Async	2	0.895	0.339	0.199	0.556
TRUE	Async	TASK ⁻	0.063	0.339	0.144	0.552
TRUE	Async	TASK ⁺	0.653	0.998	0.948	0.945
TRUE	Async	Visual	0.003	0.278	0	0.013
TRUE	Async	DMN ⁻	0.001	0.154	0.431	0.999
TRUE	Async	Post.	0	0.012	0	0.001
TRUE	Async	DMN ⁺	0.745	0.998	0.996	0.999
TRUE	Sync	state ⁰	0.907	0.998	0.326	0.561
TRUE	Sync	PH	0.996	0.998	0.525	0.999
TRUE	Sync	2	0.391	0.998	0.673	0.999
TRUE	Sync	TASK ⁻	0.391	0.998	0.13	0.945
TRUE	Sync	TASK ⁺	0.005	0.278	0.048	0.3
TRUE	Sync	Visual	0.895	0.998	0.098	0.822
TRUE	Sync	DMN ⁻	0	0.01	0.003	0.025
TRUE	Sync	Post.	0.063	0.877	0.017	0.3
TRUE	Sync	DMN ⁺	0.689	0.998	0.802	0.999

Part I: Identifying and modulating neural correlates of PH

Table 2.35 [S27] PH questionnaire

Question Number	Question Text	Sensation Assessed
1	I felt as if I was touching my back by myself	Perceived self-touch
2	I felt as if someone else was touching my back	Passivity Experiences
3	I felt as if someone was close to me or behind me	Presence Hallucination
4	I felt as if I was not controlling my movements or actions	Loss of Agency
5	I felt as if someone was in front of me	Direct control for PH phenomenology
6	I felt as if I had two bodies	General control
7	I felt anxious or stressed	Anxiety

Questionnaire presented to the participants following each condition of the PH-task performed in day 1 and in day 5. This questionnaire was rated on a Likert scale with ratings ranging from 0: "don't agree at all"; to 6: "completely agree".

Part I: Identifying and modulating neural correlates of PH

Guidance for participants regarding neurofeedback training

All participants visualized the same powerpoint slides containing pertinent information for the neurofeedback regulation task, before training started on the second day. The relevant information is summarised here:

1. Participants were first shown a slide with the various conditions of the neurofeedback training sessions, informing them on how they would know to start and stop moving (auditory cues), when the robot would switch between synchronous and asynchronous conditions, and when a certain auditory feedback – which we would later explain – would come on and off.
2. Participants were told that during the training sessions their brain activity would be monitored in real-time by the MRI scanner and that there were “types of brain activity” which were interesting for us, and that we would like them to achieve during the up-regulating condition. We explicitly told participants we could not tell them what this brain activity was related too.
3. Participants were now told that despite not knowing the goal, the auditory feedback was related to their success on regulating the desired brain activity, and that this would be given with around 5/7 seconds of delay in relation to their brain activity (haemodynamic delay + processing time).
4. Participants listened to the auditory feedback as many times as they wanted. This was repeated later inside the MRI.
5. We explain that to achieve the desire brain activity, participants should deploy a mental strategy of their choice during the up-regulation condition. Participants were told that not all strategies are necessarily explicit, and that for some the mental strategy might be something they can’t verbalise and that would be ok.
6. Participants were informed that the auditory feedback would only vary during up-regulation condition, but they would still hear it during the control condition “as to know what was the minimum level” (there were also task control purposes in doing this, that were not told to the participants).
7. We informed participants that they really should avoid any strategies during the control condition, as it could impact the perceived feedback during the interesting condition (up-regulation). At this point we suggested participants they should attempt to clear their minds.
8. Finally, we asserted strategies should be mental and gave examples of strategies that were forbidden:
 - a. Moving their feet, or any part of their body (unrelated to the robot task) – this was valid for the entire experiment
 - b. Changing the pace of stimulation of the robot between control and up-regulation
 - c. Changing breathing patterns
 - d. We recommended against focusing on the sound of the MRI
9. Participants were free to ask questions, but we guaranteed that we never gave any clues on which strategies might work best, or on our goal. Participants were blinded.

Part II: Neural correlates of hallucination severity in Parkinson's Disease

3.1 Study 3: Robot-induced hallucinations and PLS-CAPs reveal neural correlates of neurocognitive-psychiatric preservation in PD

Herberto Dhanis ^{1,2}, Jevita Potheegadoo ¹, Fosco Bernasconi ¹, Cyrille Stucker ¹, Marie Maradan ³, Laurent Jenni ¹, Joseph-André Ghika ⁴, Benoit Wicki ⁵, Pierre Burkhard ⁵, David Benninger ⁶, Judith Horvath ⁷, Paul Krack ³, Dimitri Van De Ville ², and Olaf Blanke ^{1,*}

¹ Laboratory of Cognitive Neurosciences, Neuro-X Institute, École Polytechnique Fédérale de Lausanne, CH-1202 Geneva, Switzerland

² Medical Image Processing Laboratory, Neuro-X Institute, École Polytechnique Fédérale de Lausanne, CH-1202 Geneva, Switzerland

³ University Hospital of Bern, CH-3010 Bern, Switzerland

⁴ Hospital of Sion, CH-1951 Sion, Switzerland

⁵ University Hospital of Geneva, CH-1205 Geneva, Switzerland

⁶ University Hospital of Lausanne, CH-1011 Vaud, Switzerland

⁷ Hospital de la Tour, Geneva, Switzerland

* Corresponding author

Corresponding Author

Prof. Olaf Blanke
Bertarelli Chair in Cognitive Neuroprosthetics
Neuro-X Institute & Brain Mind Institute
School of Life Sciences
Ecole Polytechnique Fédérale de Lausanne (EPFL)
Campus Biotech, H4.3
Ch. des Mines 9
CH-1202 Geneva
E-mail: olaf.blanke@epfl.ch
Tel: +41 (0)21 693 69 21

3.1.1 Abstract

Parkinson's disease (PD) is marked by different sets of hallucinations, with minor hallucinations, such as presence, passage hallucinations, and visual illusions occurring early in the disease, and structured, mostly visual hallucinations, occurring later on. Minor hallucinations typically progress to structured hallucinations, which have serious implications in disease progression and quality-of-life, leading to persistent psychosis, cognitive impairment, and higher likelihood of nursing home placement and death. Recently, a robotic device was shown induce a minor hallucination in PD patients, opening the door to its use in investigating the relationship between hallucination progression in PD, sensitivity to the procedure, cognitive impairment, and brain activity as measured with fMRI.

We recruited PD patients stratified across structured hallucinations (23 patients), minor hallucinations (11 patients), and no hallucinations (30 patients). No differences were seen across age, gender, disease duration, schooling, and severity of motor symptoms in the groups. Patients underwent neuropsychological assessments, a resting-state fMRI scanning session, and the robotic task to induce the minor hallucination. The main outcome of the neuropsychological assessment with respect to cognitive impairment was the PD Cognitive Rating Scale, whereas from the robotic task was a percentage of induction for each of three different conditions. We then used Co-Activation Pattern analysis to extract active brain networks at rest, and followed with Partial Least Squares to identify multivariate relationships between neuropsychological assessments, behavioral induction data, and activate brain networks.

Our findings show that across this cohort, patients with structured hallucinations are the most sensitive to induction, whereas those without are the least. Importantly, pairing co-activation pattern analysis and partial least squares we identified three features of brain activity which underlie, better scores in the cognitive assessment, no hallucinations, and lack of sensitivity to induction of the minor hallucination. They are: (1) less fluctuations of the visual network, (2) increased anti-correlation of the visual network with the default-mode network, paired with decreased anti-correlation of the former with the dorsal attention network, and (3) increased anti-correlation between a visuo-attentional network (intraparietal sulcus / frontal eye-field network) and the default-mode network.

We posit that these findings highlight neural correlates related to preserved cognitive capabilities and non-progression to neuropsychiatric PD pathology. Furthermore, they highlight the use of this robotic task in identifying an induction-feature relevant for the patients' neuropsychiatric status, and provide significant implications for future therapies trying to maintain these aspects of brain function as a way to stop progression to hallucinations.

Keywords: Parkinson's Disease; Hallucinations; Partial-Least Squares (PLS); Co-Activation Patterns CAPs (CAPs); Cognitive decline; Multi-variate analysis; Parkinson's Disease Dementia

3.1.2 Introduction

Hallucinations are a common non-motor symptom of Parkinson's Disease (PD), with around 70% of patients experiencing them (Fénelon et al., 2010; ffytche et al., 2017). In PD, hallucinations are mostly sub-divided in two categories: minor hallucinations (Lenka et al., 2019) and structured hallucinations (Diederich et al., 2005). Minor hallucinations include, presence hallucinations (PH), the sensation of having someone nearby when no one is present), passage hallucinations, the sensation of having something passing by you), and visual illusions which are misperceptions of visual stimuli, such as, pareidolia (Doé de Maindreville et al., 2005). Structured hallucinations, more commonly refer to visual hallucinations, although other types of hallucinations have also been reported in PD, such as auditory (Fénelon & Alves, 2010), tactile (Fénelon et al., 2002) or olfactory (Factor et al., 2014) hallucinations.

Minor hallucinations tend to occur early in the disease onset, being present in approximately 40% of patients before a diagnosis of PD is established, and potentially even before the onset of motor symptoms in some patients (Pagonabarraga et al., 2016). Their occurrence is highly predictive of the development of structured hallucinations (Lenka et al., 2019). In turn, structured hallucinations are linked to persistent psychosis (Factor et al., 2003) and a higher rate of nursing home placement and death (Hely et al., 2008; Ravina et al., 2007). An aspect of considerable interest with regards to hallucinations is cognitive impairment. Different studies have shown that worsened cognitive decline marks progression to hallucinations (Fénelon, 2000), psychosis (Lenka et al., 2017) and loss of insight (Llebaria et al., 2010). But others have indicated that the occurrence of hallucinations, might also play a role in furthering cognitive decline (Morgante et al., 2012), or found no relationship between both (Doé de Maindreville et al., 2005). The direct association between minor hallucinations and these comorbidities is less known, as outside of their progression to structured hallucinations, they have only been associated to REM sleep disorder and depression (Lenka et al., 2019). A recent study showed that minor hallucinations are associated with subjective cognitive decline, and hence could potentially be an early indicator of future cognitive decline (Bejr-kasem et al., 2021). This was further supported by a more significant grey matter loss in patients with minor hallucinations versus those without.

The brain mechanisms of minor hallucinations are also less understood than those of structured hallucinations. Visual hallucinations, part of the latter group, have been associated with higher activity of the visual associative cortex (Holroyd & Wooten, 2006), and higher activity and intra-connectivity of the default mode-network (DMN; Franciotti et al., 2015; Yao et al., 2014). Furthermore it has been hypothesized (Shine et al., 2011) and demonstrated, that crucial to visual hallucinations in PD, is a weaker recruitment of executive networks (Shine, et al., 2014) joint with an overreliance on the DMN, which increases its coupling to visual networks (typically more coupled with executive networks in patients without hallucinations; Shine et al., 2015). For minor hallucinations, to the best of our knowledge only one study investigated differences in functional connectivity, and found similar increases in intra-connectivity within the DMN, as well as between areas of the DMN and visual regions (Bejr-kasem et al., 2019). This could reflect a difficulty in recruiting patients with PD that only have minor hallucinations and have not yet progressed to structured hallucinations.

In sum, the relationship between progressive hallucinations in PD, cognitive impairment (or even decline), and underlying neural correlates is still unknown. We are in a particular position to engage with this problem as we have developed a robotic system capable of inducing PH in healthy individuals (Bernasconi et al., 2022; Blanke et al., 2014), that can in fact trigger PH with higher likelihood in PD patients with minor hallucinations versus those without (Bernasconi et al., 2021). In fact, a PH-network

Part II: Neural correlates of hallucination severity in Parkinson's Disease

associated with robot-induced PH (riPH) in healthy individuals, was already shown to have decreased connectivity in PD patients with PH versus those without, with a particular disconnection between the posterior superior temporal sulcus (pSTS) and inferior frontal gyrus (IFG) predicting up to 97% of patients with PH (Bernasconi et al., 2021).

Hence, in the current study, we set to answer these complex relationships (minor hallucinations and their progression to complex hallucinations, cognitive impairment and brain correlates), for the first time across a stratified cohort of PD patients with different degrees of hallucinations. For this, we recruited PD patients with structured and minor hallucinations (PD-SH), only minor hallucinations (PD-MH), and no hallucinations (PD-nH), as well as aged-matched controls, to take part in a study composed of two visits. A first one which included a battery of neuropsychological examinations, and a second one where patients completed a session of resting-state fMRI, followed by the robotic task to induce PH. We then studied intensity of PH induction in these patients across the hallucination spectrum and cognitive impairment. We further linked these metrics to brain network dynamics, assessed as Co-Activation Patterns' (CAPs) (Liu & Duyn, 2013) occurrences over time. Our results, show that intensity of PH-induction increases with sensorimotor conflicts, but differently depending on severity of hallucinations and cognitive impairment. Finally, studying the multivariate relationship between these clinical aspects and active brain networks, revealed that patients with intact cognition and no hallucinations, have preserved anti-correlations between an intraparietal sulcus focused network and the DMN, as well as less fluctuations of a visual network during rest, and less occurrences of a pattern with anti-correlation between the visual network and an executive network.

3.1.3 Materials and methods

3.1.3.1 Participants

All our patients were recruited at the hospitals of Geneva, Bern, Lausanne, and Sion, in Switzerland. All gave written informed consent before participant in the experiments, which were approved by the Cantonal Ethics Committee of Geneva. Patients were only included if they met specific inclusion criteria (Supplementary Table 1), and had a Montreal Cognitive Assessment (MoCA) score above 21 for PD patients or above 25 for healthy controls. A total of 76 participants took part in our experiments, consisting of 23 PD patients with structured and/or minor hallucinations (mean age: 64.44 ± 7.99 ; 16 male), 11 PD patients with only minor hallucinations (mean age: 66.75 ± 7.87 ; 7 male), 30 PD patients with no hallucinations (mean age: 66.54 ± 7.85 ; 20 male), and 12 healthy aged-matched controls (mean age: 67.14 ± 7.59 ; 7 male).

3.1.3.2 Clinical and neuropsychological evaluations

On the first visit day, every participant completed clinical and neuropsychological evaluations. For some participants this visit had to be split due to time constraints. All participants completed an in-house hallucination questionnaire designed to assess the prevalence of minor hallucinations, structured hallucinations, delusions, cognitive insight and functional disability (PD psychotic symptom scale: PD-PSS), and the PD cognitive rating scale (PD-CRS) designed to capture a considerable range of typical cognitive impairment in PD

Part II: Neural correlates of hallucination severity in Parkinson's Disease

(Pagonabarraga et al., 2008). PD patients also completed all four parts of the MDS-UPDRS (Goetz et al., 2008).

3.1.3.3 PH-induction through robotically-mediated sensorimotor conflicts

3.1.3.3.1 Sensorimotor delay dependency task

On the second visit day participants performed a robotic task designed to induce PH (Bernasconi et al., 2022), either in the center of Bern or Geneva. For all the experiments, participants were blindfolded and listened to brown noise intended to isolate them from surrounding noise. This experiment was divided into three parts.

The procedure to induce PH (Figure 3.1) was divided in 36 blocks of robot manipulation. Each block started with a single auditory cue, and finished with a double auditory cue after the participant had completed ten pokes to their back (if a participant did a front poke, but it did not had enough length to touch their back it was not counted). At this moment the participant answered verbally to the question “Did you feel as if someone was standing behind you and touching you on the back?” with “yes” or “no”. This question had been presented to the participants before the start. Two breaks of two minutes were enforced after the 12th and 24th trials. These breaks were also used to make sure the participant remembered the question.

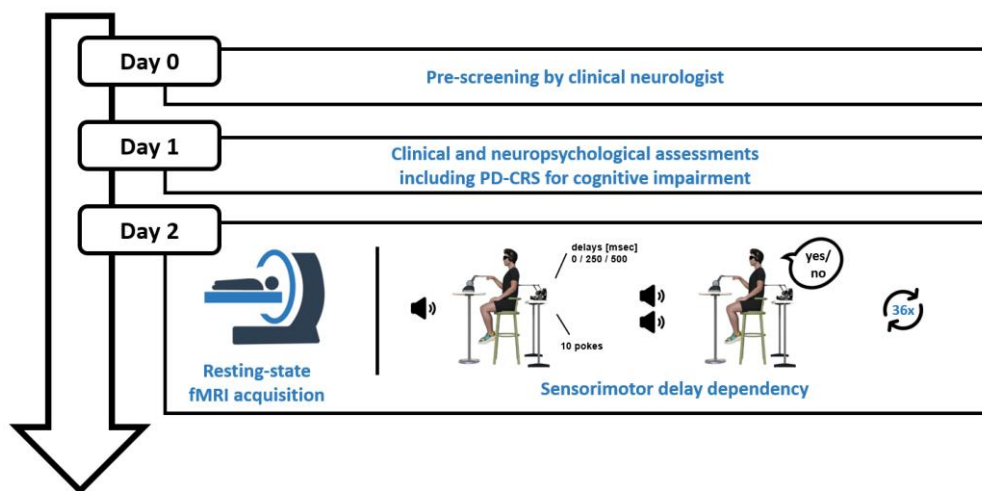


Figure 3.1 Experimental paradigm

Patients were pre-screened by their care neurologist and then came for two visits to perform our experiment. **On day 1** multiple clinical and neuropsychological assessments were done, including the PD-CRS to assess cognitive impairment. **On day 2** patients did a resting-state fMRI acquisition followed by a sensorimotor delay dependency task designed to induce riPH. Participants performed a session of 36 blocks, where after a beep, they started moving the robot (delay randomly set to 0, 250 or 500 msec). After 10 pokes were recorded, patients stopped and answered verbally to the question “Did you feel as if someone was standing behind you and touching your back?”. Answered was forced to yes or no. Patients memorised the questioned before starting and this was verified during breaks.

3.1.3.3.2 Statistical analysis

We analyzed whether disease duration, cognitive scores, delay, and hallucination grouping, modeled the intensity of riPH during the sensorimotor delay dependency task, using generalized linear mixed models (GLMM, eq. 1), given that participants gave repeated binary answers. This was performed using the function `glmm` provided with the package `lme4` (Bates et al., 2015) for R (version 4.1.3). Disease duration and cognitive scores were standardized, while delay and hallucination grouping were included as factors. A separate analysis including healthy aged-match controls was performed using only the two last predictors.

$$(eq.1) riPH\ rating \sim Delay * Group * Cogn.Score + Disease\ duration + (1|Patient)$$

We then ran type-III ANOVA's to identify predictors that significantly modulated the dependent variables. If an interaction was detected this modelled was then decomposed for one of the values of the interaction (as seen in eq. 2), and the subsequent result were correct for the false-discovery rate using the Benjamini-Hochberg method (Benjamini & Hochberg, 1995).

$$(eq.2) riPH\ rating(Group = PD_{SH}) \sim Delay * Cogn.Score + Disease\ duration + (1|Patient)$$

3.1.3.4 Imaging Data

3.1.3.4.1 Resting-state functional MRI acquisitions

Given that some of our participants had deep brain stimulation implants, not all could undergo resting-state fMRI acquisitions. For this part we counted with 16 PD-SH patients, 7 PD-MH patients, and 19 PD-nH patients. Eligible participants performed a resting-state functional MRI (fMRI) acquisition, before the behavioral experimental paradigm. The acquisition of functional images was done at the MRI facility of the Campus Biotech in Geneva (Switzerland), or at the department of neuroradiology of the university hospital of Bern (Switzerland). Both scanners were Siemens MAGNETOM Prisma, and used a 64-channel head and neck coil. The resting-state sequences used echo-planar sequences (EPI, TR = 0.7 s, TE = 30 ms, flip angle = 52°, acceleration factor of 8), with an in-plane resolution of 2-by-2 mm and a slice thickness of 2 mm (no gap, FOV = 256 mm, 72 slices). Structural images were acquired using a T1-weighted MPRAGE sequence (208 slices, FOV = 256 mm, TR = 2.2s, TE = 2.96ms).

3.1.3.4.2 fMRI Data Pre-processing

Standard CONN (Whitfield-Gabrieli & Nieto-Castanon, 2012) pipelines were used to preprocess the acquired fMRI data (these pipelines include functionalities from SPM – Wellcome Department of Cognitive Neurology, Institute of Neurology, UCL, London, UK – and from the ART toolbox – Gabrieli lab, MIT) in MATLAB (R2019b, Mathworks). Included in this preprocessing was, slice-timing correction, spatial realignment and re-slicing, removal and interpolation of frames with significant

movement (threshold at 0.5 mm), normalization to MNI space, and spatial smoothing (FWHM = 6 mm). Denoising capabilities of the CONN toolbox were used to remove confounds related to respiration and movement. Linear trends were also removed.

3.1.3.4.3 Static connectivity analysis within the PH-network

Functional connectivity was computed across the PH-network identified in previous work (Bernasconi et al., 2021). Regions of interest were established using publicly available data from said work, and transposed to the current space. These regions were the pSTS, IFG, and ventral pre-motor cortex (vPMC), all included bilaterally.

3.1.3.4.4 Co-Activation Pattern Analysis

To recover brain networks active across the patients' resting-state, we used unconstrained CAP analysis (Liu et al., 2013), using the TbCAPs toolbox (Bolton et al., 2020). While typical CAP analysis requires the selection of timepoints where a seed region presents with high activity, unconstrained CAPs applies a clustering method to all acquired time-frames using k-means. Here we selected $K = 10$, based on consensus clustering (Stefano Monti et al., 2003) performed at patient-level, and then assessed the results across patients. Once the clustering is run each time point is assigned to a certain centroid, and representative centroids (CAPs) can be generated by averaging time frames with the same assigned CAP. To improve the quality of our results we discard any time-frame with frame-wise displacement above 0.5 mm (Power et al., 2012), and 10% of the frames with minimal activation. Based on these results we then extracted the occurrence probability of each CAP, by computing the total number of occurrences over the total number of frames.

3.1.3.4.5 Partial Least Squares CAPs

In order to assess any potential multivariate relationship across the collected behavioral data, cognitive assessments, and CAPs occurrences, we applied correlation-PLS analysis to these data. In sum, a cross-covariance matrix is computed from the clinical/behavioral data and CAP occurrences matrixes over patients. Group contrasts and interactions with behavioral data were included in the clinical/behavioral matrix, hence covariance maximization was done across all patients. Singular value decomposition of this matrix then leads to a number of latent components (LC), that are composed by sets of weights of CAP occurrences and sets of clinical/behavioral weights. Imaging and behavioral loadings were then obtained by correlating each, respectively to the original CAP occurrences and behavioral data. These represent how much each variable within the sets contributes to the brain-clinical/behavior correlation.

The significance of the LCs was assessed doing 1000 permutation tests, and the stability of the loadings through a bootstrapping procedure. To guarantee that we did not bias the analysis we performed the permutation and bootstrapping within groups, as to respect differences observed at the level of the clinical/behavioral data across groups.

3.1.4 Results

3.1.4.1 Demographic, clinical and neuropsychological assessments

Across group, patients did not show any difference in gender, age, years of schooling, disease duration, MoCA scores, and UPDRS-III scores (all: $P > 0.05$). However, PD-SH patients were taking a significantly higher daily L-dopa equivalent dosage ($P < 0.001$). The PD-CRS showed differences across groups for the fronto-cortical scores ($P < 0.001$), but no for the cortical scores ($P > 0.05$).

3.1.4.2 Sensorimotor delay dependency task

Results from the sensorimotor delay dependency task show that overall riPH increased with delay ($P < 0.001$), and with cognitive impairment ($P < 0.05$). A triple-interaction was detected between delay, group, and cognitive impairment ($P < 0.0001$). Decomposition of the model across groups allowed to quantify this effect (all results were corrected for multiple comparisons). The sensitivity to riPH of SH-PH, was significantly higher than that of the other groups ($P < 0.05$), and overall all groups continued showing a significant increase of riPH with delay (all $P < 0.001$; Figure 3.2A). This effect of delay was further modulated by cognitive impairment for PD-MH and PD-nH (2-way interactions; both: $P < 0.01$, Figure 3.2B). Cognitive impairment was found to significantly increase the sensitivity to riPH for both delay (250 and 500 msec) in the PD-nH group (both $P > 0.01$). No impact was seen for PD-MH.

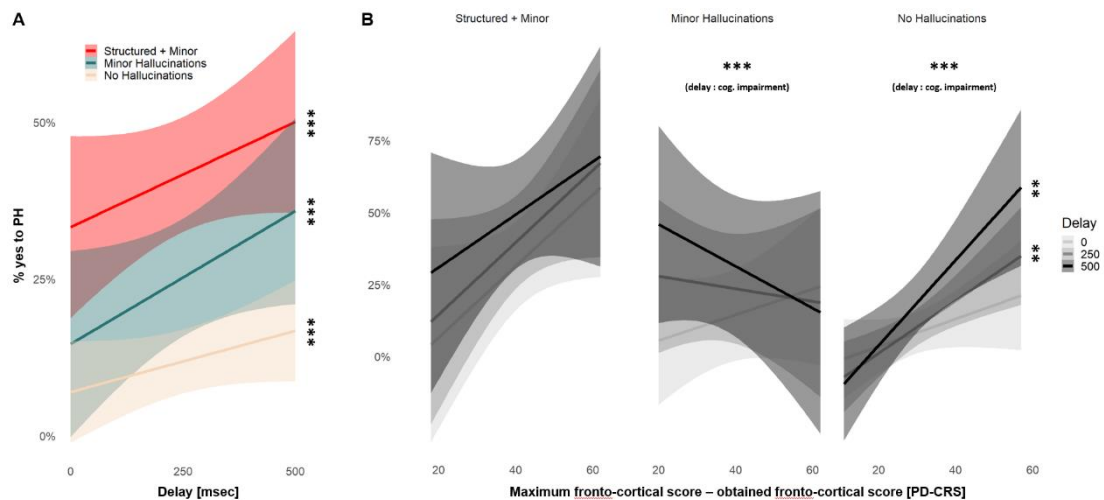


Figure 3.2 Results of the sensorimotor delay dependency task

(A) Percentages of yes to experience riPH are shown in function of delay for PD-SH (red), PD-MH (turquoise), and PD-nH (beige). LMM analysis showed a significant increase in riPH with increasing delay (all $P < 0.001$). (B) Percentages of yes to riPH are shown for each patient group, in function of increasing cognitive impairment computed as the maximum value of the fronto-cortical score from the PD-CRS assessment minus the actual obtained score. An interaction for both PD-MH and PD-nH was observed between delay and cognitive impairment (both $P < 0.001$). Results for PD-SH are shown but no interaction was detected, meaning the effect of cognitive impairment was the same across every delay. Subsequent analysis did not identify any significant difference on the PD-MH group, but identified a significant effect which increased riPH with an increase in cognitive impairment (both $P < 0.01$).

3.1.4.3 CAPs analysis

CAP analysis revealed how different brain patterns came about across the different groups of patients. With this we identified 10 CAPs, described here in order of prevalence, that captured all the major networks described in PD with relevance for hallucinations and for cognitive impairment. CAP 1 (Figure 3.3A) identified an activated visual network (VN) paired with a deactivated dorsal attention network (DAN). CAPs 2 and 3 identified the VN in and activated and deactivated form, respectively (Figures 3.3B, 3.3C). CAP 4 (Figure 3.3D) identified a deactivated DMN paired with activations over the supramarginal gyri (SMG). CAP 5 the DAN paired with a deactivated VN (Figure 3.3E). CAP 6 identified an activated VN paired with deactivations over the posterior DMN (Figure 3.3F). CAPs 7 and 8 identified an intraparietal sulcus/frontal eye-field network (IPS-FEF-N) and DMN which presented in deactivated/activated and activated/deactivated forms, respectively (Figures 3.3F, 3.3H). CAP 9 identified an activated posterior DMN, paired with a deactivated VN (Figure 3.3I). And finally CAP 10 identified an activated DMN paired with deactivations over the SMG and frontal regions (Figure 3.3J).

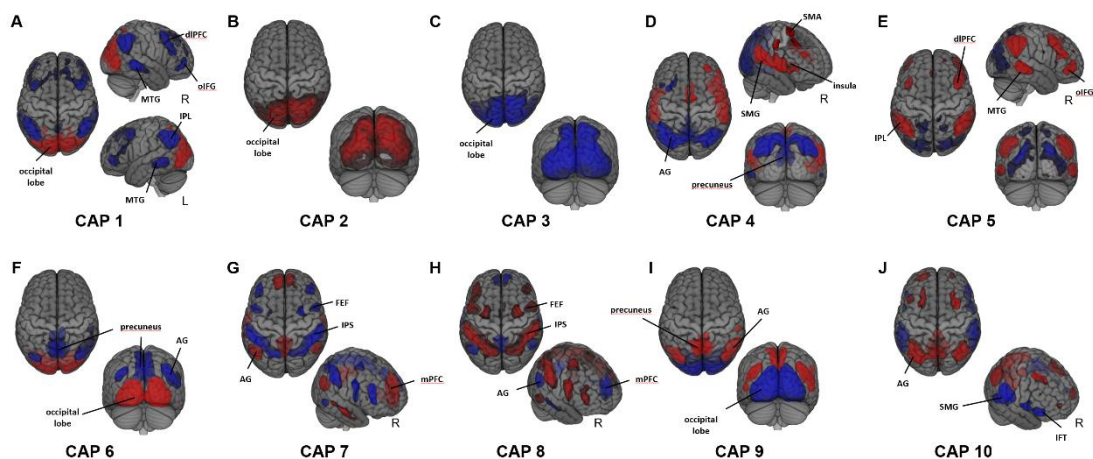


Figure 3.3 Identified Co-Activation Patterns

Z-scored maps of the CAPs are shown, with activations (red) and deactivations (blue) above 1.5 standard deviations. **(A)** CAP1: activated visual network (VN) and deactivated central executive network (CEN), with the latter including the dorso-lateral prefrontal cortex (dlPFC), middle temporal gyrus (MTG), orbital inferior frontal gyrus (olFG), and inferior parietal lobule (angular gyrus + supramarginal gyrus; AG, SMG). **(B)** Activated VN. **(C)** Deactivated VN. **(D)** Activations over regions responsible for salience mapping including the insula and the anterior cingulate cortex (ACC), as well as the supramarginal gyrus (SMG), and supplemental motor area (SMA), paired with deactivations over the posterior default mode network (DMN), including, AG and precuneus. **(E)** Activated CEN paired with deactivations over the cuneus. **(F)** Activated VN paired with deactivations over the posterior DMN. **(G)** Activated DMN, including precuneus, AG, and medial pre frontal cortex (mPFC), paired with deactivations over the dorsal attentional network (DAN), which includes the intraparietal sulcus (IPS) and frontal eye field (FEF). **(H)** Activated DAN paired with deactivated DMN. **(I)** Activated DMN and deactivated VN. **(J)** Activated DMN with deactivations over the SMG and inferior frontal triangularis (IFT). Note that: CAPs 1 and 5, 2 and 3, 4 and 10, 6 and 9, 7 and 8, all represent reversed patterns, respectively.

3.1.4.4 PLS-CAPs

After computing the occurrences for each of the identified CAPs, we used PLS to identify multivariate relationships between the imaging data and both the clinical data and behavioral data regarding PH-induction. This revealed one significant latent component (LC; $P < 0.05$). With regards to imaging, this LC was sensitive to a decrease in occurrences of CAPs 1, 2, and 3, and an increase in occurrences of CAPs 6, 7, and 8 (Figure 3.4A). Regarding clinical-behavioral data, the LC was associated with the contrast PD-nH > PD-MH/PD-SH, less disease duration, less cognitive impairment, and less mean sensitivity to rIPH (albeit paired with a positive slope across delays; Figure 3.4B).

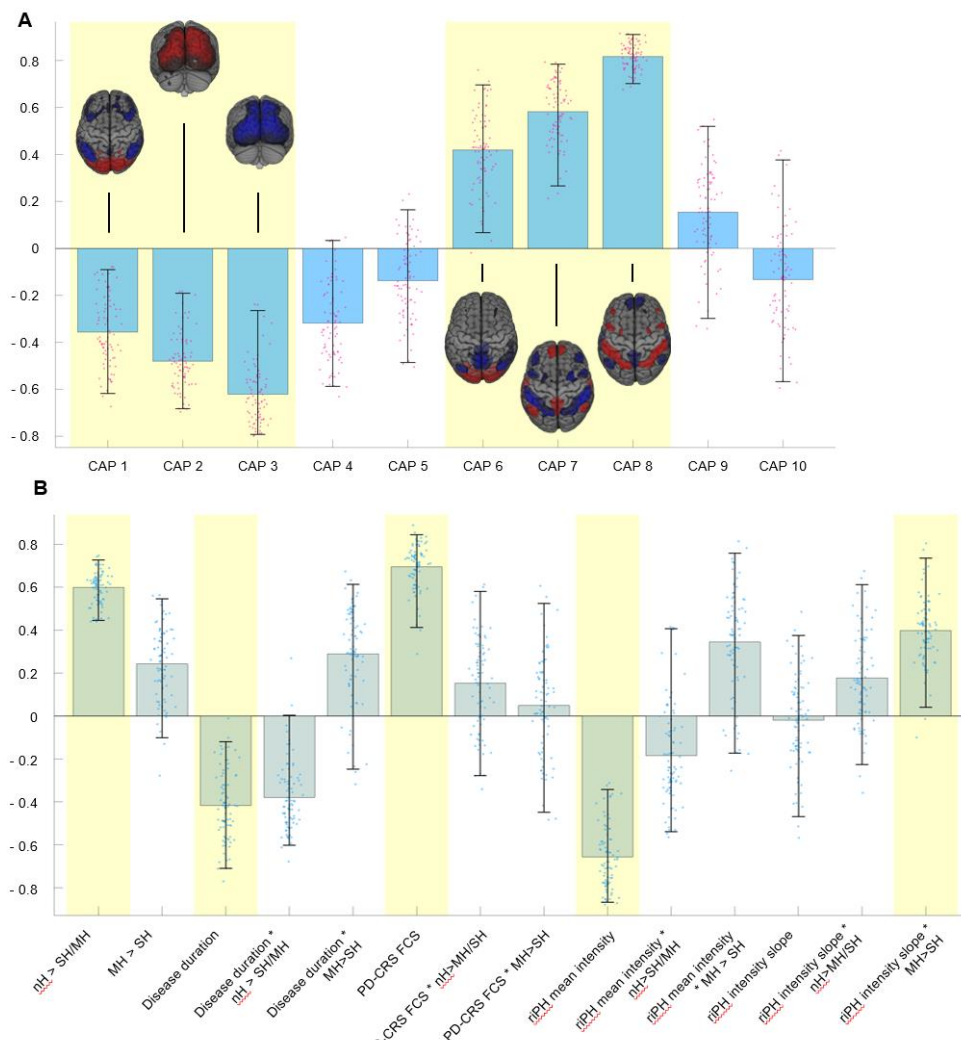


Figure 3.4 PLS-CAPs analysis

(A) Imaging loadings from the PLS CAPs, showing CAPs with significantly decreased occurrences (1, 2 and 3), and significantly increased occurrences (6, 7, and 8), highlighted with a yellow background. CAP 1 shows activated visual areas and deactivations over the DAN. CAP 2 and CAP 3 represent the same pattern with activations or deactivations over visual areas, respectively. CAP 6 includes activations in visual areas paired with deactivations over the DMN. CAP 7 and CAP 8 represent the same pattern: activations/deactivations on the DMN, paired with deactivations/activations over the IPS/FEF network. **(B)** Design loadings from the PLS-CAPs analysis. Yellow background highlights significant variables.

3.1.5 Discussion

We combined state-of-the-art robotics to induce hallucinations, neuropsychological examinations, and resting-state fMRI, in a cohort of PD patients stratified by hallucination severity, to reveal brain correlates associated with hallucination symptomatology and cognitive impairment. Our results, were able to show hallucination severity and cognitive impairment increase the sensitivity of the patients to riPH, and importantly, we were able to identify a set of brain patterns which marked preserved cognition, lack of hallucinations, and unsensitivity to our robotic procedure to induce PH. With this it could be possible that our robotic procedure distinguishes between a worse form of the disease marked by hallucinations and cognitive impairment.

3.1.5.1 Sensorimotor delay dependency stratifies hallucination sensitivity through cognitive impairment

We used an established robotic procedure (Bernasconi et al., 2022) to measure our patients sensitivity to a specific minor hallucination. In line with previous work (Bernasconi et al., 2021), our results showed that the robotic system triggered riPH with higher probability in patients with PH (i.e., PD-MH) versus patients without PH (i.e., PD-nH). New in our study was the inclusion of patients with structured and minor hallucinations. This group revealed to be the most sensitive to riPH on average, followed by PD-MH, and by PD-nH which had the lowest sensitivity. In addition, delay increased the sensitivity to riPH in all groups.

Analyzing the effect of cognitive impairment in our task, across groups and delays (triple interaction: group, delay, disease duration), revealed that increased cognitive impairment lead to an increase in the sensitivity to riPH across the 250 and 500 msec delay conditions for the PD-nH group. In practice this means cognitive impairment is increasing the sensitivity with delay. PD-nH participants are seen to have approximately the same riPH sensitivity for all delays at minimal cognitive impairment, but as cognitive impairment increases so does the distance between riPH sensitivity for the different delays (being highest for 500 msec). Regarding PD-MH, a significant interaction between delay and cognitive impairment was also seen, with the data suggesting the opposite effect seen for PD-nH (sensitivity decreases with cognitive impairment, and all delays end up having the same approximate sensitivity). However, these results did not reach significance after decomposition of the interaction. We do note, this group had a lower sample size than the other groups. No interaction between cognitive impairment and delay was seen for PD-SH.

Being the second study of its kind to analyze PH-induction in PD patients, our results greatly extend previous work, which only focused on patients with PH or no PH, and provides significant evidence of a continuum of sensitivity to riPH across hallucination severity in these patients. Albeit in a different hallucination modality, work in visual hallucinations in PD, has also revealed that PD patients with visual hallucinations are more prone to visual illusions than PD patients without visual hallucinations (Shine et al., 2012). Finally, the fact that cognitive impairment affects our task differently between groups, is an indicator that PH-induction via our robotic task could be of use to further understand clinical progression of these patients.

3.1.5.2 Neural correlates of absent hallucinatory pathology, lack of sensitivity to riPH and preserved cognition in patients with PD

By pairing resting-fMRI in a stratified cohort of PD patients, with robotics for the induction of a clinical relevant hallucination and with neuropsychological assessments of cognitive decline, we identified three characteristics of brain dynamics which correlate with an absence of symptomatic hallucinations, lack of sensitivity to riPH and a more preserved cognition (also referred to as PD-nH preserved). These characteristics are increased antagonistic activity between the IPS-FEF-N and DMN, an increase in antagonistic activity between the VN and DMN together with a decrease between the former and the DAN, and decreased fluctuations of the VN.

3.1.5.2.1 Relationship between the VN and both DAN and DMN

Through PLS-CAPs we found two different brain patterns indicating that the VN had increased antagonistic activity with the DMN, paired with decreased with the DAN, in PD-nH with preserved cognition and lack of sensitivity to riPH, who were early in the disease course. We hypothesize that the observation of this pattern, reveals that a dysfunction of over-reliance on the DMN that has been hypothesized to underlie visual hallucinations in PD (Shine et al., 2011; Shine, et al., 2014), has not occurred for our PD-nH patients that fit the preserved category. Furthermore recent studies have shown that when PD patients with visual hallucinations incur in visual misperceptions during a bistable perception task (Shine et al., 2012), network connectivity between the VN and the DAN decreases, and rather increases between the VN and DMN (Shine et al., 2015). Another study looking at mind-wandering in patients with visual hallucinations in PD, also found that these patients had increased coupling between the VN and the DMN (Walpola et al., 2020). These are accompanied by observations in different studies of intra and inter hyperconnectivity of the DMN in PD patients with visual hallucinations versus patients without hallucinations (Franciotti et al., 2015; Yao et al., 2014). Our current findings suggest that the opposite pattern marks PD patients that do not have hallucinations, are cognitively preserved, and not sensitive to our protocol.

3.1.5.2.2 Decreased activity of the Visual Network

As mentioned, before the VN has been found to have dysfunctional interactions with key networks including the DAN and DMN (Shine et al., 2015). Here aside from observing that the VN remained decoupled from the DAN and DMN, we further observed that PD-nH preserved patients exhibited less fluctuations of the VN. While this specific finding has not yet been reported in PD with hallucinations, a recent study using multimodal imaging found significant dysfunctions on the visual system of PD patients with visual hallucinations (versus those without), which included functional and structural abnormalities (Diez-Cirarda et al., 2023). Although not directly tested here, we consider it possible that the decrease in VN fluctuations might indicate a preserved connectome, with certain states being preferred over others (with in this case the VN not being preferred during resting-state). This proposition tries to conciliate our findings with recent observations in PD patients

with visual hallucinations who show increased ease in transitioning across different brain states, while patients with no hallucinations were able to stay longer in a state and transition less to others (Zarkali et al., 2022).

3.1.5.2.3 Retained visuo-spatial attentional network function

Two CAPs in our analysis identified strong antagonistic activity between the IPS-FEF-N and the DMN, in association with PD-nH preserved patients. While the DMN as mentioned above is seen with dysfunctional connectivity in PD patients with visual hallucinations, the role of the IPS-FEF-N has been mostly seen in relation to visual attention (see for example Offen et al., 2010). In particular, one study in PD has linked the IPS-FEF-N with attention dysfunction, and showed that PD patients engage this network more than healthy individuals during an attentional task, and also show decreased connectivity between the IPS and the precuneus (a node of the DMN) (Boord et al., 2017). Here, we found that preserved antagonistic activity between the DMN and IPS-FEF-N were associated with preserved cognition amongst others. This would be in line with increased functional connectivity (anti-correlation), and with major theories which propose that attentional networks should preserve anti-correlations to the DMN, in both PD (Shine et al., 2011), but also other conditions that include psychopathology (Menon, 2011).

3.1.5.3 Limitations

Our results advance the understanding of brain dysfunction related to hallucinations in PD, with a particular focus on correlates of retained cognition and absence of hallucinations. However, they are not without limitations. While here we have identified correlates of preserved functional aspects in patients with PD, we cannot know if these patients will develop hallucinations, in a way a longitudinal study would. We have nonetheless tried to address this by including patients across a significant range of disease duration for all group, including PD-nH. Another limitation is the fact that we are not yet capable of triggering riPH inside the MRI-scanner in the same manner done for healthy individuals in other studies (Bernasconi et al., 2021; Dhanis et al., 2022), which would have allowed to study the direct neural correlates of hallucinations in the different groups of patients. Finally, we note that in our cohort PD-SH patients were seen to be on a higher daily L-dopa equivalent dosage that was higher than other groups.

3.1.5.4 Conclusions

The results of our study provide the first direct evidence of both, increased sensitivity of induced hallucinations with increasing hallucination severity, and neural correlates of preserved neuropsychiatric and cognitive function. Crucially, this was done across a sample of PD patients stratified based on hallucination severity, in a manner that is rare in literature (PD-nH, PD-MH, and PD-SH). Our results indicate that preserved functional characteristics of the VN, DMN, and two attentional networks (VAN and IPS-FEF-N), are significantly correlated with preserved cognition, lack of hallucinations and lack of sensitivity to having them triggered by an established robotic procedure. Given that the present robot was already used in an fMRI neurofeedback study that achieved control in modulation of an hallucination network (with

Part II: Neural correlates of hallucination severity in Parkinson's Disease

subsequent modulation of the hallucination; Dhanis et al., 2022), these findings could have implications for potential therapies trying to preserve these functional traits of brain activity in an attempt to prevent neuropsychiatric progression in PD.

3.1.6 References

- Andreasen, N. C. (1984). Scale for the Assessment of Positive Symptoms [Data set]. American Psychological Association. <https://doi.org/10.1037/t48377-000>
- Bates, D., Mächler, M., Bolker, B., & Walker, S. (2015). Fitting Linear Mixed-Effects Models Using lme4. *Journal of Statistical Software*, 67(1). <https://doi.org/10.18637/jss.v067.i01>
- Beck, A. (2004). A new instrument for measuring insight: The Beck Cognitive Insight Scale. *Schizophrenia Research*, 68(2–3), 319–329. [https://doi.org/10.1016/S0920-9964\(03\)00189-0](https://doi.org/10.1016/S0920-9964(03)00189-0)
- Bejr-kasem, H., Pagonabarraga, J., Martínez-Horta, S., Sampedro, F., Marín-Lahoz, J., Horta-Barba, A., Aracil-Bolaños, I., Pérez-Pérez, J., Ángeles Botí, M., Campolongo, A., Izquierdo, C., Pascual-Sedano, B., Gómez-Ansón, B., & Kulisevsky, J. (2019). Disruption of the default mode network and its intrinsic functional connectivity underlies minor hallucinations in Parkinson's disease. *Movement Disorders*, 34(1), 78–86. <https://doi.org/10.1002/mds.27557>
- Bejr-kasem, H., Sampedro, F., Marín-Lahoz, J., Martínez-Horta, S., Pagonabarraga, J., & Kulisevsky, J. (2021). Minor hallucinations reflect early gray matter loss and predict subjective cognitive decline in Parkinson's disease. *European Journal of Neurology*, 28(2), 438–447. <https://doi.org/10.1111/ene.14576>
- Benjamini, Y., & Hochberg, Y. (1995). Controlling the false discovery rate A practical and powerful approach to multiple testing. 57(1), 289–300.
- Bernasconi, F., Blondiaux, E., Potheegadoo, J., Stripeikyte, G., Pagonabarraga, J., Bejr-Kasem, H., Bassolino, M., Akselrod, M., Martinez-Horta, S., Sampedro, F., Hara, M., Horvath, J., Franza, M., Konik, S., Bereau, M., Ghika, J.-A., Burkhard, P. R., Van De Ville, D., Faivre, N., ... Blanke, O. (2021). Robot-induced hallucinations in Parkinson's disease depend on altered sensorimotor processing in fronto-temporal network. *Science Translational Medicine*, 13(591), eabc8362. <https://doi.org/10.1126/scitranslmed.abc8362>
- Bernasconi, F., Blondiaux, E., Rognini, G., Dhanis, H., Jenni, L., Potheegadoo, J., Hara, M., & Blanke, O. (2022). Neuroscience robotics for controlled induction and real-time assessment of hallucinations. *Nature Protocols*. <https://doi.org/10.1038/s41596-022-00737-z>
- Blanke, O., Pozeg, P., Hara, M., Heydrich, L., Serino, A., Yamamoto, A., Higuchi, T., Salomon, R., Seeck, M., Landis, T., Arzy, S., Herbelin, B., Bleuler, H., & Rognini, G. (2014). Neurological and Robot-Controlled Induction of an Apparition. *Current Biology*, 24(22), 2681–2686. <https://doi.org/10.1016/j.cub.2014.09.049>
- Bolton, T. A. W., Tuleasca, C., Wotruba, D., Rey, G., Dhanis, H., Gauthier, B., Delavari, F., Morgenroth, E., Gaviria, J., Blondiaux, E., Smigielski, L., & Van De Ville, D. (2020). TbCAPs: A toolbox for co-activation pattern analysis. *NeuroImage*, 211, 116621. <https://doi.org/10.1016/j.neuroimage.2020.116621>
- Boord, P., Madhyastha, T. M., Askren, M. K., & Grabowski, T. J. (2017). Executive attention networks show altered relationship with default mode network in PD. *NeuroImage: Clinical*, 13, 1–8. <https://doi.org/10.1016/j.nicl.2016.11.004>

Part II: Neural correlates of hallucination severity in Parkinson's Disease

Cardebat, D., Démonet, J. F., Viallard, G., Faure, S., Puel, M., & Celsis, P. (1996). Brain Functional Profiles in Formal and Semantic Fluency Tasks: A SPECT Study in Normals. *Brain and Language*, 52(2), 305–313. <https://doi.org/10.1006/brln.1996.0013>

Christensen, R. H. B. (2022). ordinal—Regression Models for Ordinal Data (R package version 2022.11-16.). <https://CRAN.R-project.org/package=ordinal>

Dhanis, H., Blondiaux, E., Bolton, T., Faivre, N., Rognini, G., Van De Ville, D., & Blanke, O. (2022). Robotically-induced hallucination triggers subtle changes in brain network transitions. *NeuroImage*, 248, 118862. <https://doi.org/10.1016/j.neuroimage.2021.118862>

Diederich, N. J., Goetz, C. G., & Stebbins, G. T. (2005). Repeated visual hallucinations in Parkinson's disease as disturbed external/internal perceptions: Focused review and a new integrative model. *Movement Disorders*, 20(2), 130–140. <https://doi.org/10.1002/mds.20308>

Diez-Cirarda, M., Cabrera-Zubizarreta, A., Murueta-Goyena, A., Strafella, A. P., Del Pino, R., Acera, M., Lucas-Jiménez, O., Ibarretxe-Bilbao, N., Tijero, B., Gómez-Esteban, J. C., & Gabilondo, I. (2023). Multimodal visual system analysis as a biomarker of visual hallucinations in Parkinson's disease. *Journal of Neurology*, 270(1), 519–529. <https://doi.org/10.1007/s00415-022-11427-x>

Doé de Maindreville, A., Fénelon, G., & Mahieux, F. (2005). Hallucinations in Parkinson's disease: A follow-up study. *Movement Disorders*, 20(2), 212–217. <https://doi.org/10.1002/mds.20263>

Factor, S. A., Feustel, P. J., Friedman, J. H., Comella, C. L., Goetz, C. G., Kurlan, R., Parsa, M., Pfeiffer, R., & the Parkinson Study Group. (2003). Longitudinal outcome of Parkinson's disease patients with psychosis. *Neurology*, 60(11), 1756–1761. <https://doi.org/10.1212/01.WNL.0000068010.82167.CF>

Factor, S. A., Scullin, M. K., Sollinger, A. B., Land, J. O., Wood-Siverio, C., Zanders, L., Freeman, A., Bliwise, D. L., McDonald, W. M., & Goldstein, F. C. (2014). Cognitive correlates of hallucinations and delusions in Parkinson's disease. *Journal of the Neurological Sciences*, 347(1–2), 316–321. <https://doi.org/10.1016/j.jns.2014.10.033>

Fénelon, G. (2000). Hallucinations in Parkinson's disease: Prevalence, phenomenology and risk factors. *Brain*, 123(4), 733–745. <https://doi.org/10.1093/brain/123.4.733>

Fénelon, G., & Alves, G. (2010). Epidemiology of psychosis in Parkinson's disease. *Journal of the Neurological Sciences*, 289(1–2), 12–17. <https://doi.org/10.1016/j.jns.2009.08.014>

Fénelon, G., Soulas, T., Zenasni, F., & de Langavant, L. C. (2010). The changing face of Parkinson's disease-associated psychosis: A cross-sectional study based on the new NINDS-NIMH criteria. *Movement Disorders*, 25(6), 763–766. <https://doi.org/10.1002/mds.22839>

ffytche, D. H., Creese, B., Politis, M., Chaudhuri, K. R., Weintraub, D., Ballard, C., & Aarsland, D. (2017). The psychosis spectrum in Parkinson disease. *Nature Reviews Neurology*, 13(2), 81–95. <https://doi.org/10.1038/nrneurol.2016.200>

Franciotti, R., Delli Pizzi, S., Perfetti, B., Tartaro, A., Bonanni, L., Thomas, A., Weis, L., Biundo, R., Antonini, A., & Onofri, M. (2015). Default mode network links to visual hallucinations: A comparison between Parkinson's disease and multiple system atrophy: Default Mode Network Links to Visual Hallucinations. *Movement Disorders*, 30(9), 1237–1247. <https://doi.org/10.1002/mds.26285>

Part II: Neural correlates of hallucination severity in Parkinson's Disease

Fénelon, G., Thobois, S., Bonnet, A.-M., Broussolle, E., & Tison, F. (2002). Tactile hallucinations in Parkinson's disease. *Journal of Neurology*, 249(12), 1699–1703. <https://doi.org/10.1007/s00415-002-0908-9>

Goetz, C. G., Tilley, B. C., Shaftman, S. R., Stebbins, G. T., Fahn, S., Martinez-Martin, P., Poewe, W., Sampaio, C., Stern, M. B., Dodel, R., Dubois, B., Holloway, R., Jankovic, J., Kulisevsky, J., Lang, A. E., Lees, A., Leurgans, S., LeWitt, P. A., Nyenhuis, D., ... LaPelle, N. (2008). Movement Disorder Society-sponsored revision of the Unified Parkinson's Disease Rating Scale (MDS-UPDRS): Scale presentation and clinimetric testing results: MDS-UPDRS: Clinimetric Assessment. *Movement Disorders*, 23(15), 2129–2170. <https://doi.org/10.1002/mds.22340>

Haggard, P., & Chambon, V. (2012). Sense of agency. *Current Biology*, 22(10), R390–R392. <https://doi.org/10.1016/j.cub.2012.02.040>

Hely, M. A., Reid, W. G. J., Adena, M. A., Halliday, G. M., & Morris, J. G. L. (2008). The Sydney multicenter study of Parkinson's disease: The inevitability of dementia at 20 years: Twenty Year Sydney Parkinson's Study. *Movement Disorders*, 23(6), 837–844. <https://doi.org/10.1002/mds.21956>

Holroyd, S., & Wooten, G. F. (2006). Preliminary fMRI evidence of visual system dysfunction in Parkinson's Disease patients with visual hallucinations. *The Journal of Neuropsychiatry and Clinical Neurosciences*, 18(3), 402–404.

Hughes, M. E., Waite, L. J., Hawkey, L. C., & Cacioppo, J. T. (2004). A Short Scale for Measuring Loneliness in Large Surveys: Results From Two Population-Based Studies. *Research on Aging*, 26(6), 655–672. <https://doi.org/10.1177/0164027504268574>

Lenka, A., George, L., Arumugham, S. S., Hegde, S., Reddy, V., Kamble, N., Yadav, R., & Pal, P. K. (2017). Predictors of onset of psychosis in patients with Parkinson's disease: Who gets it early? *Parkinsonism & Related Disorders*, 44, 91–94. <https://doi.org/10.1016/j.parkreldis.2017.09.015>

Lenka, A., Pagonabarraga, J., Pal, P. K., Bejr-Kasem, H., & Kulisevsky, J. (2019). Minor hallucinations in Parkinson disease: A subtle symptom with major clinical implications. *Neurology*, 93(6), 259–266. <https://doi.org/10.1212/WNL.00000000000007913>

Liu, X., Chang, C., & Duyn, J. H. (2013). Decomposition of spontaneous brain activity into distinct fMRI co-activation patterns. *Frontiers in Systems Neuroscience*, 7. <https://doi.org/10.3389/fnsys.2013.00101>

Liu, X., & Duyn, J. H. (2013). Time-varying functional network information extracted from brief instances of spontaneous brain activity. *Proceedings of the National Academy of Sciences*, 110(11), 4392–4397. <https://doi.org/10.1073/pnas.1216856110>

Llebaria, G., Pagonabarraga, J., Martínez-Corral, M., García-Sánchez, C., Pascual-Sedano, B., Gironell, A., & Kulisevsky, J. (2010). Neuropsychological correlates of mild to severe hallucinations in Parkinson's disease: Neuropsychology of VH in Parkinson's Disease. *Movement Disorders*, 25(16), 2785–2791. <https://doi.org/10.1002/mds.23411>

Menon, V. (2011). Large-scale brain networks and psychopathology: A unifying triple network model. *Trends in Cognitive Sciences*, 15(10), 483–506. <https://doi.org/10.1016/j.tics.2011.08.003>

Part II: Neural correlates of hallucination severity in Parkinson's Disease

Mlakar, J., Jensterle, J., & Frith, C. D. (1994). Central monitoring deficiency and schizophrenic symptoms. *Psychological Medicine*, 24(3), 557–564. <https://doi.org/10.1017/S0033291700027719>

Morgante, L., Colosimo, C., Antonini, A., Marconi, R., Meco, G., Pederzoli, M., Pontieri, F. E., Cicarelli, G., Abbruzzese, G., Zappulla, S., Ramat, S., Manfredi, M., Bottacchi, E., Abrignani, M., Berardelli, A., Cozzolino, A., Paradiso, C., De Gaspari, D., Morgante, F., ... on behalf of the PRIAMO Study Group. (2012). Psychosis associated to Parkinson's disease in the early stages: Relevance of cognitive decline and depression. *Journal of Neurology, Neurosurgery & Psychiatry*, 83(1), 76–82. <https://doi.org/10.1136/jnnp-2011-300043>

Offen, S., Gardner, J. L., Schluppeck, D., & Heeger, D. J. (2010). Differential roles for frontal eye fields (FEFs) and intraparietal sulcus (IPS) in visual working memory and visual attention. *Journal of Vision*, 10(11), 28–28. <https://doi.org/10.1167/10.11.28>

Pagonabarraga, J., Kulisevsky, J., Llebaria, G., García-Sánchez, C., Pascual-Sedano, B., & Gironell, A. (2008). Parkinson's disease-cognitive rating scale: A new cognitive scale specific for Parkinson's disease: Cognitive Rating Scale for PD. *Movement Disorders*, 23(7), 998–1005. <https://doi.org/10.1002/mds.22007>

Pagonabarraga, J., Martinez-Horta, S., Fernández de Bobadilla, R., Pérez, J., Ribosa-Nogué, R., Marín, J., Pascual-Sedano, B., García, C., Gironell, A., & Kulisevsky, J. (2016). Minor hallucinations occur in drug-naïve Parkinson's disease patients, even from the premotor phase: Minor Hallucinations in Untreated PD Patients. *Movement Disorders*, 31(1), 45–52. <https://doi.org/10.1002/mds.26432>

Power, J. D., Barnes, K. A., Snyder, A. Z., Schlaggar, B. L., & Petersen, S. E. (2012). Spurious but systematic correlations in functional connectivity MRI networks arise from subject motion. *NeuroImage*, 59(3), 2142–2154. <https://doi.org/10.1016/j.neuroimage.2011.10.018>

Radakovic, R., & Abrahams, S. (2014). Developing a new apathy measurement scale: Dimensional Apathy Scale. *Psychiatry Research*, 219(3), 658–663. <https://doi.org/10.1016/j.psychres.2014.06.010>

Ravina, B., Marder, K., Fernandez, H. H., Friedman, J. H., McDonald, W., Murphy, D., Aarsland, D., Babcock, D., Cummings, J., Endicott, J., Factor, S., Galpern, W., Lees, A., Marsh, L., Stacy, M., Gwinn-Hardy, K., Voon, V., & Goetz, C. (2007). Diagnostic criteria for psychosis in Parkinson's disease: Report of an NINDS, NIMH work group. *Movement Disorders*, 22(8), 1061–1068. <https://doi.org/10.1002/mds.21382>

Reitan, R. M., & Wolfson, D. (1985). The Halstead-Reitan neuropsychological test battery: Theory and clinical interpretation (Vol. 4). Reitan Neuropsychology.

Ruff, R. M., Niemann, H., & Wylie, T. (1994). Figural Fluency: Differential Impairment in Patients with Left Versus Right Frontal Lobe Lesions. *Archives of Clinical Neuropsychology: The Official Journal of the National Academy of Neuropsychologists*, 9(1), 41–55.

Schmitt, E., Krack, P., Castrioto, A., Klinger, H., Bichon, A., Lhommée, E., Pelissier, P., Fraix, V., Thobois, S., Moro, E., & Martinez-Martin, P. (2018). The Neuropsychiatric Fluctuations Scale for Parkinson's Disease: A Pilot Study: the neuropsychiatric fluctuations scale for Parkinson's disease: a pilot study. *Movement Disorders Clinical Practice*, 5(3), 265–272. <https://doi.org/10.1002/mdc3.12607>

Part II: Neural correlates of hallucination severity in Parkinson's Disease

Shine, J. M., Halliday, G. H., Carlos, M., Naismith, S. L., & Lewis, S. J. G. (2012). Investigating visual misperceptions in Parkinson's disease: A novel behavioral paradigm: Visual Perceptions in PD. *Movement Disorders*, 27(4), 500–505. <https://doi.org/10.1002/mds.24900>

Shine, J. M., Halliday, G. M., Gilat, M., Matar, E., Bolitho, S. J., Carlos, M., Naismith, S. L., & Lewis, S. J. G. (2014). The role of dysfunctional attentional control networks in visual misperceptions in Parkinson's disease: Visual Misperceptions in Parkinson's Disease. *Human Brain Mapping*, 35(5), 2206–2219. <https://doi.org/10.1002/hbm.22321>

Shine, J. M., Halliday, G. M., Naismith, S. L., & Lewis, S. J. G. (2011). Visual misperceptions and hallucinations in Parkinson's disease: Dysfunction of attentional control networks?: Attentional Control Networks in PD Hallucinations. *Movement Disorders*, 26(12), 2154–2159. <https://doi.org/10.1002/mds.23896>

Shine, J. M., Muller, A. J., O'Callaghan, C., Hornberger, M., Halliday, G. M., & Lewis, S. J. (2015). Abnormal connectivity between the default mode and the visual system underlies the manifestation of visual hallucinations in Parkinson's disease: A task-based fMRI study. *Npj Parkinson's Disease*, 1(1), 15003. <https://doi.org/10.1038/npjparkd.2015.3>

Shine, J. M., O'Callaghan, C., Halliday, G. M., & Lewis, S. J. G. (2014). Tricks of the mind: Visual hallucinations as disorders of attention. *Progress in Neurobiology*, 116, 58–65. <https://doi.org/10.1016/j.pneurobio.2014.01.004>

Stefano Monti, Pablo Tamayo, Jill Mesirov, & Todd Golub. (2003). Consensus Clustering: A Resampling-Based Method for Class Discovery and Visualization of Gene Expression Microarray Data. *Machine Learning*, 52, 91–118.

Stiasny-Kolster, K., Mayer, G., Schäfer, S., Möller, J. C., Heinzel-Gutenbrunner, M., & Oertel, W. H. (2007). The REM sleep behavior disorder screening questionnaire-A new diagnostic instrument. *Movement Disorders*, 22(16), 2386–2393. <https://doi.org/10.1002/mds.21740>

Walpola, I. C., Muller, A. J., Hall, J. M., Andrews-Hanna, J. R., Irish, M., Lewis, S. J. G., Shine, J. M., & O'Callaghan, C. (2020). Mind-wandering in Parkinson's disease hallucinations reflects primary visual and default network coupling. *Cortex*, 125, 233–245. <https://doi.org/10.1016/j.cortex.2019.12.023>

Wechsler, D. (2008). Wechsler adult intelligence scale—Fourth Edition (WAIS-IV). San Antonio, TX: NCS Pearson, 22(498), 1.

Weintraub, D., Hoops, S., Shea, J. A., Lyons, K. E., Pahwa, R., Driver-Dunckley, E. D., Adler, C. H., Potenza, M. N., Miyasaki, J., Siderowf, A. D., Duda, J. E., Hurtig, H. I., Colcher, A., Horn, S. S., Stern, M. B., & Voon, V. (2009). Validation of the questionnaire for impulsive-compulsive disorders in Parkinson's disease: ICD Questionnaire for Parkinson's Disease. *Movement Disorders*, 24(10), 1461–1467. <https://doi.org/10.1002/mds.22571>

Whitfield-Gabrieli, S., & Nieto-Castanon, A. (2012). Conn: A Functional Connectivity Toolbox for Correlated and Anticorrelated Brain Networks. *Brain Connectivity*, 2(3), 125–141. <https://doi.org/10.1089/brain.2012.0073>

Yao, N., Shek-Kwan Chang, R., Cheung, C., Pang, S., Lau, K. K., Suckling, J., Rowe, J. B., Yu, K., Ka-Fung Mak, H., Chua, S., Ho, S. Leong, & McAlonan, G. M. (2014). The default mode network is disrupted in

Part II: Neural correlates of hallucination severity in Parkinson's Disease

parkinson's disease with visual hallucinations. *Human Brain Mapping*, 35(11), 5658–5666. <https://doi.org/10.1002/hbm.22577>

Zarkali, A., Luppi, A. I., Stamatakis, E. A., Reeves, S., McColgan, P., Leyland, L.-A., Lees, A. J., & Weil, R. S. (2022). Changes in dynamic transitions between integrated and segregated states underlie visual hallucinations in Parkinson's disease. *Communications Biology*, 5(1), 928. <https://doi.org/10.1038/s42003-022-03903-x>

Zigmond, A. S., & Snaith, R. P. (1983). The Hospital Anxiety and Depression Scale. *Acta Psychiatrica Scandinavica*, 67(6), 361–370. <https://doi.org/10.1111/j.1600-0447.1983.tb09716.x>

Part III: Quantifying and optimizing PH induction

4.1 Mega-Analysis of Presence Hallucination Induction Experiments Using Robotically Mediated Sensorimotor Conflicts

Herberto Dhanis ^{1,2}, Jevita Potheegadoo ^{1,2}, Nathan Faivre ³, Fosco Bernasconi ^{1,2}, and Olaf Blanke ^{1,2,4,*}

¹ Centre for Neuroprosthetics, Faculty of Life Sciences, Swiss Federal Institute of Technology (Ecole Polytechnique Fédérale de Lausanne, EPFL), Geneva, Switzerland

² Brain Mind Institute, Faculty of Life Sciences, Swiss Federal Institute of Technology (Ecole Polytechnique Fédérale de Lausanne), Lausanne, Switzerland

³ University Grenoble Alpes, University Savoie Mont Blanc, CNRS, LPNC, Grenoble, France

⁴ Department of Neurology, University Hospital of Geneva, Geneva, Switzerland

* Corresponding author

Corresponding Author

Prof. Olaf Blanke
Bertarelli Chair in Cognitive Neuroprosthetics
Neuro-X Institute & Brain Mind Institute
School of Life Sciences
Ecole Polytechnique Fédérale de Lausanne (EPFL)
Campus Biotech, H4.3
Ch. des Mines 9
CH-1202 Geneva
E-mail: olaf.blanke@epfl.ch
Tel: +41 (0)21 693 69 21

4.1.1 Abstract

Hallucinations are significant symptoms in many psychiatric and neurodegenerative diseases, that can furthermore indicate disease progression or a worse form of the disease. Despite their relevance studying hallucinations in controlled laboratorial conditions remains complicated due to their erratic nature, and lack of experimental protocols that can induce clinically relevant hallucinations similar to what patients experience in daily life. Recently, a robotics approach has been successful in inducing a clinically relevant hallucination, presence hallucination (PH), in healthy individuals and patients afflicted by it. Since its development, several studies have used this setup to investigate the mechanisms of PH-induction, its neural correlates through an MR-compatible version of the system, and how PH-induction affects different cognitive functions, by pairing the robotic setup with specific experimental paradigms. With the growing number of studies using this approach and the ongoing translation to clinical populations, here, we have conducted a Bayesian meta-analysis over 26 experiments (580 participants) that used this robotic system in healthy individuals, to establish the effect size of induction and of other experimental parameters, which varied across experiments. In doing so, we established that this system can induce PH with an effect size of +0.33 (95% HDI: 0.12, 0.55). Based on our assessments we also make recommendations on the use of different experimental parameters, control conditions, and identified relationships between increased induction and schizotypy in the general population.

4.1.2 Introduction

Hallucinations are complex and heterogeneous phenomena during which an individual has an aberrant perceptual experience, in the absence of any corresponding external stimuli. As a symptomatic phenomenon, hallucinations have been described in psychiatric conditions, such as schizophrenia (Whitfield-Gabrieli and Ford, 2012), in neurological disorders such as stroke or infarctions affecting sensory cortices (de Haan et al., 2007; Flint et al., 2005), and neurodegenerative diseases like Lewy body dementia (Nagahama et al., 2007) and Parkinson's disease (Ffytche H et al., 2017). The clinical relevance of such hallucinations is significant. For example, it is one of the most prevalent clinical features of schizophrenia (Bilder et al., 2000), and it might be an early maker of severe forms of Parkinson's disease, given that the occurrence of so called minor hallucinations (presence and passage hallucinations, visual illusions) is linked to progression towards more complex hallucinations, loss of insight, delusions, and dementia (Lenka et al., 2019; Ravina et al., 2007). In particular, for the latter, significant research has focused on the Presence Hallucination (PH) - the convincing and realistic feeling of having someone close by when no one is actually there (Brugger et al., 1997) - as it is one of the most common hallucination in Parkinson's disease (Fénelon et al., 2011). For many patients it is also the one with earliest onset, around 42% before diagnosis, and around 15% before motor symptom onset (Pagonabarraga et al., 2016). PH is however not specific to Parkinson's disease. It is also frequent in a variety of other conditions such as schizophrenia (46%, Llorca et al., 2016) and Lewy body dementia (20%, Nagahama et al., 2010), and to a lesser degree has also been described in stroke (Blanke et al., 2003), epilepsy (Arzy et al., 2006), and in healthy individuals undergoing extreme circumstances such as high-altitude mountaineers (Brugger et al., 1999; Messner, 2016), and shipwreck survivors (Geiger, 2009). PH is most often experienced on the back or to the side of the afflicted individual, sometimes sharing the same body traits, and mimicking movements, postures and actions (Blanke et al., 2008, 2003; Brugger et al., 1997, 1996; Potheegadoo et al., 2022). Despite its high degree of perceived realism, PH has no clear established sensory component, as opposed to formed visual or auditory-verbal hallucinations (Brugger et al., 1997; Picard, 2010). This symptomatology paired with the fact that some patients feel a sense of "like-mindedness" regarding the presence, led several neurologists to group PH with alterations of the body schema and altered sensorimotor processing (Blanke et al., 2008; Blanke et al., 2014). Consequently, PH is considered a misperceived double of the self, arising due to impaired sensorimotor processing (Brugger et al., 1996; Case et al., 2019), a hypothesis that has gained further support due to emerging clinical evidence of neurological dysfunctions of different etiologies over sensorimotor processing hubs (in stroke: Blanke et al., 2003; in invasive electrical stimulation of an epileptic patient undergoing pre-surgical evaluations: Arzy et al., 2006; in neurological patients: Blanke et al., 2014).

However, despite these clinical data and the clinical relevance of PH (e.g. in Parkinson's disease), accurate descriptions of its mechanisms remain elusive due to a variety of factors. Hallucinations tend to be unpredictable, occurring during daily life activities most often far from clinical settings. The prejudices associated with hallucinatory experiences are known to refrain patients (and healthy individuals) from reporting them (Pagonabarraga et al., 2014; Wood et al., 2015). In Parkinson's disease this has been further exacerbated by the absence of a systematic screening by clinicians (Chan and Rossor, 2002). Furthermore, even if hallucinations were reported, their description and interpretation are likely to be confounded by both patient's and researcher's biases (Berrios and Marková, 2002; Nisbett and Wilson, 1977). Finally, studying hallucinations in specific patient cohorts suffers from a potential confound of disease, given that disease-specific traits might interact with the state processes behind hallucinations (Bernasconi, et al., 2022; Kühn and Gallinat, 2012). Inducing and

Part III: Quantifying and optimizing PH induction

studying the same hallucinations in healthy individuals would be a way to circumvent some of these issues, and different methods have indeed been developed to do so, although with limited success. For instance, flicker-induced (Allefeld et al., 2011) and Ganzfeld-induced hallucinations (Wackermann et al., 2008a) generate hallucinations experimentally. However, in both cases the induced hallucinations vary largely in phenomenological content, are incomparable to symptomatic hallucinations observed in medical practice in most cases, and lack experimental control conditions. Drug-induced hallucinations, although associated with a variety of hallucinatory states, tend to be modality unspecific and are riddled with confounding factors, ranging from behavioral changes to significant alterations and/or impairments of consciousness (Baggott et al., 2010; Timmermann et al., 2018). In that sense, there is a lack for experimental methods and procedures capable of inducing specific clinically-relevant hallucinations, in the laboratory, in real time and with appropriate control conditions.

Recently, a new procedure has been proposed to tackle this problem. By integrating clinical observations of impaired sensorimotor processing in PH (Arzy et al., 2006; Blanke et al., 2014, 2003) with advances in cognitive neuroscience and robotics, Blanke and colleagues (2014) developed a new protocol capable of inducing PH in fully controlled experimental conditions (Bernasconi, et al. in press). In this setup, designed to expose individuals to sensorimotor conflicts, participants move with their dominant hand a robot placed in front of them (front-robot), while a back-robot provides tactile feedback to their backs by replicating the participant's front movements. When operating this system with the two robots in synchrony participants do not experience PH, but rather report control sensations, such as self-touch. It is only when a temporal conflict is introduced between the participants' movements and the tactile feedback on the back (asynchronous condition), thus increasing the perceived sensorimotor conflicts, that participants report a robot-induced PH (riPH). Although the primary goal of this setup is the induction of PH in this asynchronous condition, participants also report induced passivity experiences (PE; robot-induced PE: riPE), the sensation that someone else is the agent producing the touch on their back (Mlakar et al., 1994), and changes in Sense of Agency (SoA), the sensation of being the agent behind one's actions (Haggard and Chambon, 2012). The setup initially developed for healthy individuals (Blanke et al., 2014), has been extended to patients with psychiatric conditions (Salomon et al., 2020), and to patients with Parkinson's disease (Bernasconi et al., 2021). The latter study highlighted the procedure's ability to induce PH in Parkinson's disease patients, and revealed a subgroup of patients with PH as a clinical symptom in daily life whom both linked the riPH to the symptomatic one, and showed higher sensitivity to riPH than patients without such symptomatic PH. Further developments of this setup were extended to study the impact of PH induction in different perceptual and cognitive processes, such as auditory-verbal processing, auditory-verbal hallucinations (Salomon et al., 2020), auditory misperceptions (Orepic et al., 2021), fluency and memory tasks (Serino et al., 2021), and metacognition (Faivre et al., 2020). Lastly, this robotic procedure has now, also been extended to the MRI (Bernasconi et al., 2021), allowing the investigation of the neural correlates of riPH during induction and the description of the involved brain mechanisms of PH (Bernasconi et al., 2021; Dhanis et al., 2022) and other associated sensations such as riPE (Dhanis et al., 2022).

We now consider it paramount to properly quantify the effect that different experimental parameters used in various paradigms had on the induction of riPH, and advise on the optimal setup. To do so, we considered that Bayesian estimation would be the most adequate tool, as it can provide estimations of effects with associated distributional information (Kruschke and Liddell, 2018). Furthermore, we expected this large pooling of data to be informative, for parameters that varied between

experiments, and for parameters that varied within experiment but might have had small effect sizes not detectable in smaller populations. Thus, we carried out a Bayesian mega-analysis on data obtained from 26 experiments (580 healthy participants), that used the same robotic system to induce PH in healthy individuals. Our main goals were, to determine the effect size of the main experimental modulator (asynchrony) in PH induction, and to quantify the effect of the distinct experimental parameters that were investigated in the different experimental protocols of each study for the induction of riPH and associated illusions (i.e., riPE and SoA). Given the considerable amount of data available, we also intended to estimate the distribution underlying the ordinal scale used to rate riPH, which is a direct indication of the population-level bias when rating the riPH. Such an estimation can bring significant power to future studies, as they will be able to run models with the appropriate parameters for the underlying distribution. This will alleviate future studies from estimating the underlying distribution from relatively small amounts of data, or from making the likely erroneous assumption that the underlying distribution is metric (Liddell and Kruschke, 2018). As a secondary goal, we investigated the relationship between riPH, riPE and changes in SoA, the latter two related to PH in terms of phenomenology (Bernasconi et al., 2021; Blanke et al., 2014), but with distinct neural mechanisms (Dhanis et al., 2022).

4.1.3 Results

4.1.3.1 Included studies, characteristics and assessed parameters

Included in this mega-analysis are 26 experiments conducted in 14 different studies, mostly in the Laboratory of Cognitive Neurosciences, EPFL, Switzerland. Of these, 12 experiments constitute already published data, 11 experiments are in submission process, and 3 other experiments constitute pilot data (previously unpublished, released with the present mega-analysis – see supplemental material – in an effort to counter the publication bias that affects meta-analysis: Dickersin, 2005; Sharpe, 1997). All included studies here can be seen in Table 1. This comprised a total population of 580 healthy individuals (291 males) with a mean age of 24.1 years old (SD: 4.9 years old; range: 7 to 56 years old).

We used a Bayesian model to assess if several experimental factors changed questionnaire responses, given in the form of a Likert scale (0 to 6) and after completion of the PH induction task (see methods). Our model did not assume any informative priors and included flexible thresholds for the distribution underlying the ordinal response. Modelled parameters included the main riPH modulator across experiments: asynchrony (temporal delay between hand movement and tactile feedback on the back); experimental parameters that varied within or between experiments: body location (participants either received tactile feedback on the back or on the hand; hand condition included in 78 participants), robot-type (93 participants with MRI robot, 487 with standing robot), previous exposure (defined as participants having had some period of robot manipulation before assessment: 243 participants with, 337 without), duration of the robot manipulation period (varied across experiments), order (randomized across all participants), cognitive load (defined as participants having a task concomitant with riPH assessment: 137 participants with, 443 without), virtual back (defined as a resistance produced by the front-robot as to mimic participants touching a surface in front: 108 participants with, 472 without); and demographic

and trait characteristics: age, handedness score (Edinburgh Handedness Inventory: EDI; Oldfield, 1971), and delusional ideation score (Peter's Delusional Inventory: PDI; Peters et al., 2004). Interactions were also assessed between asynchrony and all the other parameters (see methods for descriptive formulas and random effect structure). Sex at birth was included as a covariate of no-interest. Finally, parameters were considered to modulate riPH if their mode and high-density interval (HDI) was completely outside of the region of practical equivalence (ROPE) for the null value (Kruschke and Liddell, 2018a), established here as an interval of 10% of the population standard deviation (SD) away from zero (null value). This interval was established as it represents very small and hence neglectable effect sizes. Conversely, we considered no modulation of riPH if the mode was inside the ROPE, or that this analysis provided inconclusive results if the mode was outside the ROPE but its HDI overlapped with the ROPE.

4.1.3.2 Effect of asynchrony of the sensorimotor stimulation on riPH

The main modulator of riPH was the asynchronous (vs. synchronous) manipulation of the robotic device, in all 26 experiments included in this mega-analysis. The posterior estimate for the effect of asynchrony on riPH varied across these different experiments (Figure 4.1), suggesting that experimental changes lead to differences in riPH. Globally, the asynchronous sensorimotor stimulation resulted in higher ratings than the synchronous condition (asynchrony: +0.33 SD, 95% HDI: 0.12, 0.55, oHR: 0%; Figure 4.2A), as assessed through our model, suggesting this condition is more likely to elicit a PH as proposed in the original study (Blanke, et al. 2014).

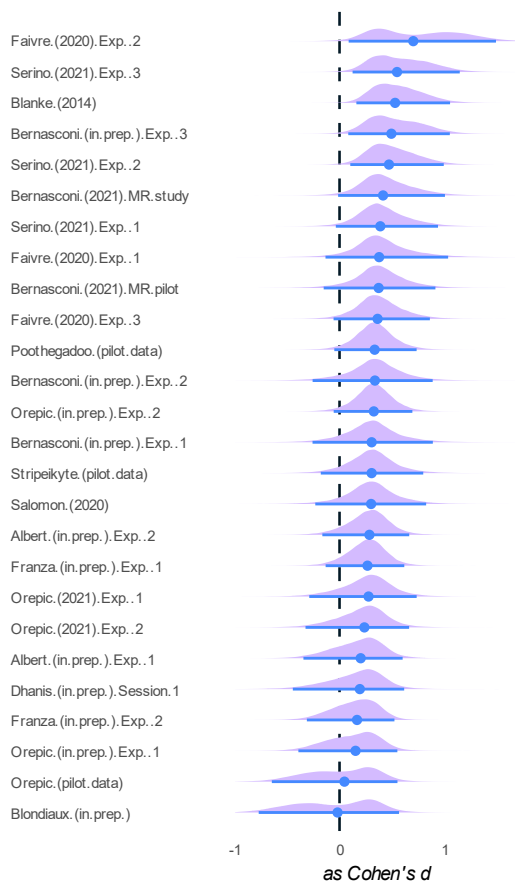


Figure 4.1 Estimates for the effect of asynchrony across all experiments

Forest plot highlighting the different estimates for the asynchrony effect across multiple experiments. The experiments are identified on the left with the code used in table 1. The estimates are shown on the right as Cohen's d , with the distribution of the estimate in purple, the blue circle indicating the mode, and the blue line the 95% high density interval of the estimate.

4.1.3.3 Effect of different experimental parameters on PH induction

In line with these changes across experiments, we assessed if different experimental parameters used across different studies modulated riPH.

Both body-location and robot-type modulated the intensity of riPH. For body-location, ratings were higher when feedback was provided to the back of participants versus the hand (mode for the effect of moving stimulation from the back to the hand: -0.66 SD, 95% HDI: [-0.95, -0.26] , oHR = 0%; figure 4.2B). While this suggested that the hand condition lead to a more residual riPH in general, we were unable to conclude on a decisive effect for the interaction between asynchrony and body location (mode: 0.01, 95% HDI: [-0.44, 0.46], oHR = 35.96%; figure 4.2C) given the overlap between HDI and ROPE. Hence, while a more residual PH might still be induced when using peripheral stimulation, it is likely that the asynchrony effect remains preserved in most cases. Regarding robot-type, we found that the MR-compatible setup showed a negative effect when compared to the classical standing robot (mode: -0.65, 95% HDI: [-1.22; -0.11], oHR = 0%; figure 4.2D). Given that no interaction was observed with asynchrony (mode: 0.01, 95% HDI: [-0.40, 0.51], oHR = 34.04%; figure 4.2E), this might, as for body-location, imply that the asynchrony effect is still present, and hence the MR-compatible robot might simply lead to a more residual, albeit still present, riPH.

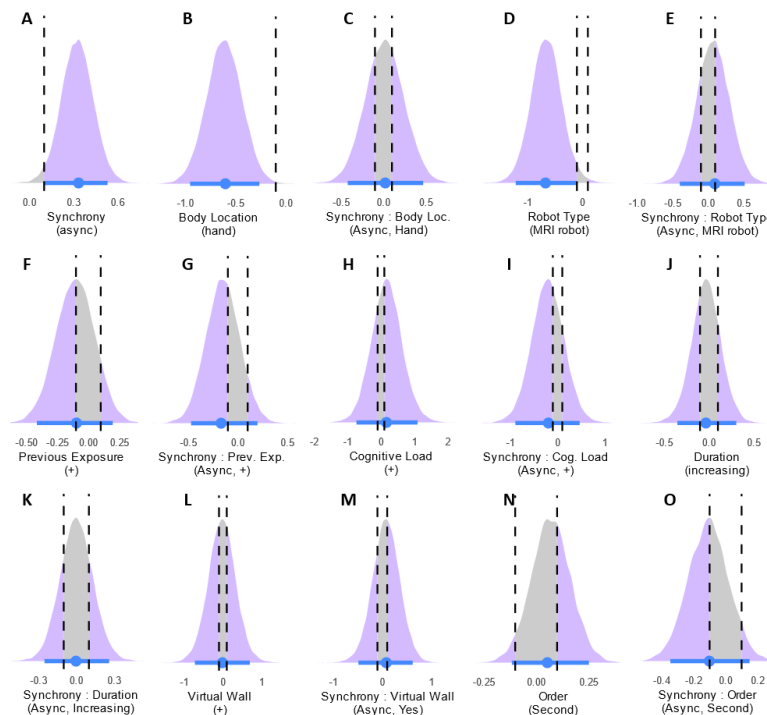


Figure 4.2 Posterior estimates for the experimental parameters modulating riPH

Posterior estimates for the various experimental parameters or interactions between parameters and synchrony are shown. Below and in parenthesis is indicated the direction of the effect. E.g. Synchrony (async) represents the estimate of the async effect as compared to sync. In purple is the distribution of the estimate with the blue circle indicating the mode and the blue line the high density interval (HDI). Vertical dashed black bars, indicate the limits of the region of practical equivalence (ROPE). The distribution of the estimate in show in grey when it overlaps with the ROPE.

Part III: Quantifying and optimizing PH induction

For the remainder of the experimental parameters the estimation of their effects did not allow a clear conclusion on the modulation of riPH, due to the overlap of the respective posterior estimates' HDIs and the ROPE for the null value: previous exposure (mode for the effect of adding previous exposure: -0.10, 95% HDI: [-0.40, 0.20], oHR = 41.75%; figure 4.2F); cognitive load (mode for the effect of adding cognitive load: 0.17, 95% HDI: [-0.70, 1.07], oHR = 17.35%; figure 4.2H); duration of the experiment (mode for increasing duration effect: -0.04, 95% HDI: [-0.36, 0.29], oHR = 49.59%; figure 4.2J); order of conditions (mode for the effect of a condition coming in second: 0.05, 95% HDI: [-0.12, 0.25], oHR = 62.14%; figure 4.2K); and virtual back (mode for the effect of adding a virtual back: 0.02, 95% HDI: [-0.76, 0.64], oHR = 24.04%; figure 4.2M). This was also the case for the interaction of these parameters with asynchrony: previous exposure (mode: -0.18, 95% HDI: [-0.4, 0.19], oHR = 32.64%; figure 4.2G), cognitive load (mode: -0.16, 95% HDI: [-0.54, 0.20], oHR = 19.82%; figure 4.2I), duration (mode: 0, 95% HDI: [-0.25, 0.25], oHR = 59.64%, figure 4.2K), virtual wall (mode: 0.07, 95% HDI: [-0.47, 0.63], oHR = 29.32%, figure 4.2M) and order (mode: -0.10, 95% HDI: [-0.36, 0.15], oHR = 45.01%; figure 4.2O).

4.1.3.4 Effect of demographic and trait characteristics on PH induction

In this analysis we have also investigated the effects of age, handedness and delusional ideation on the ratings of riPH. Handedness and delusional ideation were not available for all the participants (EHI: 12 experiments, 289 participants; PDI: 8 experiments, 178 participants), and special precautions were taken when modelling this data (see methods). Sex at birth was included as a covariate of no interest.

Only delusional ideations were shown to modulate the ratings of riPH, with higher PDI scores leading to higher riPH ratings (mode: 0.28; 95% HDI: 0.12; 0.43, oHR = 0%; figure 4.3D). This suggested that individuals with delusional ideations are generally more sensitive to riPH, independently of condition. The interaction between asynchrony and PDI did not lead to conclusive results (mode: -0.01, 95% HDI: [-0.24, 0.20], oHR = 65.55%, figure 4.3F), hence it is possible that the asynchrony effect is preserved in similar fashion across individuals with different levels of delusional ideations. However, we do note that decomposition of the general PDI effect, for each ordinal value of the scale revealed that participants are more likely

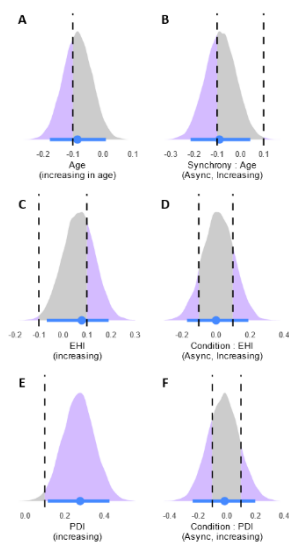


Figure 4.3 Posterior estimate for age and trait characteristics of participants
Estimates for are shown for characteristics associated with the participants and for their interactions with synchrony. In purple is the distribution of the estimate with the blue circle indicating the mode and the blue line the high density interval (HDI). Vertical dashed black bars, indicate the limits of the region of practical equivalence (ROPE). The distribution of the estimate in show in grey when it overlaps with the ROPE.

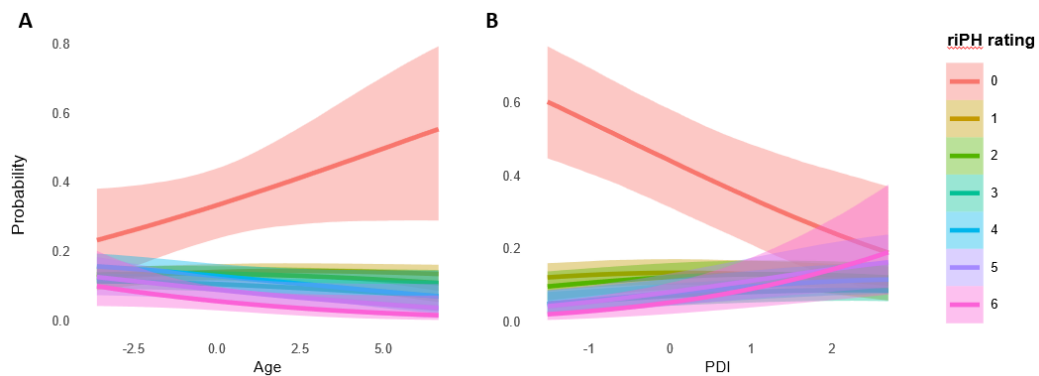


Figure 4.4 Estimation of the effect of age and PDI across the ordinal scale

Decomposition of the beta estimates for the effects of age (A) and PDI (B) are shown for each ordinal value of the scale. On the x axis we show the standardize change in the regression coefficient, and on the y axis how the standardized change modulates the probability of rating each value of the scale. (A) For age a clear distinction is identified between the item “0” and the remaining ones. With increasing age, participants have a higher probability of rating “0”. This is not observed for the other items which mostly remain stable across age, with the higher items “5” and “6” potentially experiencing a slight decrease in probability with increase in age. (B) Regarding delusional ideations assessed through PDI, a clear distinction is observed for the item “0” which has a high probability of being rated for participants with very low PDI, and then sharply decreasing in probability with increasing PDI. Item “6” so a slight opposite pattern with the probability increasing for individuals with higher delusional ideations. The remaining items remain stable across PDI scores.

to rate 6 the higher their PDI scores were, and are more likely to rate 0, the lower their PDI scores were (figure 4.4A). In particular, the latter, might imply that changes in PDI scores modulate general sensitivity to riPH.

Our analysis did not produce conclusive results for the remaining characteristics and respective interactions with asynchrony, given the posterior estimates’ HDIs and their overlap with the ROPE. Age: mode for increasing age was -0.09 SD, 95% HDI: [-0.17, 0.01], oHR = 63.42% (figure 4.3B); handedness: mode for increasing EHI: 0.08, 95% HDI: [-0.07, 0.19], oHR = 70.03% (figure 4.3C). Interactions with asynchrony: age, mode = -0.09, 95% HDI = [-0.17, 0.01], oHR = 60.06%; handedness: mode = 0, 95% HDI = [-0.17, 0.19], oHR = 75.41%. However, we note that for age, a decomposition of the effect for each ordinal value of the scale, suggests that younger participants are less likely to rate the extreme lower scale rating “0”, when compared with older participants (figure 4.4B).

4.1.3.5 Distribution underlying the riPH rating scale

Given the large amount of data analyzed here, we intended to adequately estimate the distribution underlying 7-point Likert scale used to rate riPH. This was done through the inclusion of flexible threshold parameters for the intercept values of the scale, and visualization of the underlying scale achieved through a generative model which only included these 6 parameters. Figures 4.5A and 4.5B show, respectively, the real distribution of the data, and the underlying distribution of the scale once the data is generated with this limited intercept model. It is exceedingly more likely for participants to rate “0” (mode = 32.56%, 95% HDI = [24.20, 43.97]) than the remaining values of the scale (figure 4.5C-D).

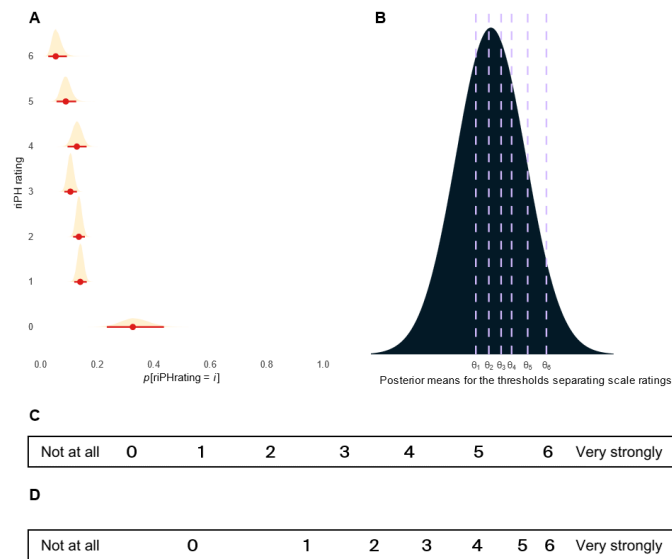


Figure 4.5 Characteristics of the ordinal rating scale used for riPH

(A) Probability of rating each of the items of the rating scale after accounting for the effect of every experimental, demographic, and trait characteristic. In blue is the mode of the probability, with the blue line being the 95% HDI associated to this probability. In purple is the associated distribution. **(B)** Normal distribution with the posterior means of the thresholds separating the values of the ordinal scale. **(C)** Visual representation of the ordinal scale used to rate riPH, as seen visually by the participants. **(D)** The same scale transformed to reflect the inner biases of participants when rating the scale, according to the posterior means described in (B).

4.1.3.6 Effects on passivity experiences, sense of agency and control questions

Here we report on the effects of all experimental parameters, demographic and trait characteristics, on the two other induced sensations: PE and changes in SoA.

The mode of the effect of asynchrony in the manipulation of the robotic device was estimated to be higher for robot-induced PE (riPE; mode = 0.46, 95% HDI = [0.25; 0.67], oHR = 0%) than for riPH, albeit overlap of the HDI with the HDI from riPH. For SoA the estimate for the mode was also higher (mode = 0.87, 95% HDI = [0.52, 1.22], oHR = 0%) than for riPH, however here, without overlap of HDIs. We note, that whereas riPE induction was investigated for the same number of experiments as riPH, changes in SoA were only studied in 9 experiments (189 healthy participants).

Regarding the effects of the other experimental parameters (type of robot, duration, etc), all had either, no effect on the ratings of riPE and SoA, or their effect was inconclusive due to overlaps between the estimates' HDIs and the ROPE for the null value.

With respect to the demographic and trait characteristics, only delusional ideation was seen to modulate both riPE and SoA. In both cases, the effects mimicked what was observed for riPH, albeit less pronounced. For changes in SoA, although the same global effect was observed, the decomposition for the different ordinal values of the scale, showed only a decreased probability to rate "0" with higher PDI scores (the opposite relationship for higher PDI scores, and the rating "6", was not observed).

4.1.3.7 Relationship between the inductions of PH, PE, and changes in SoA

Intensity of riPH was reported alongside with the riPE in all studies, while changes in SoA were reported in nine studies. Taking individual participant's data across all studies, we assessed if the intensity of riPH was modulated by both sensitivity to riPE and to changes in SoA, and if the intensity of riPE was modulated by both sensitivity to riPH and again changes in SoA.

To do so we adapted our models to include sensitivity predictors in a similar fashion to a mediation analysis, and hypothesized a relationship between asynchrony, sensitivity to changes in SoA, sensitivity to riPH/riPE, and an increased modulation of the intensity or riPE/riPH due to these (figure 4.6A). This was modelled in stages given data limitations (see methods; figure 4.6B-D).

Mediation analysis for the intensity of riPH Sensitivity to changes in SoA did not conclusively modulated riPH (mode = 0.31, 95% HDI = [-0.22, 0.77], oHR = 27.71%, figure 4.7A). Sensitivity to riPE was shown to increase the intensity of riPH (mode = 0.47, 95% HDI = [0.18, 0.73], oHR = 0%; figure 4.7B), and in a similar fashion being sensitive to both changes in SoA and riPE, on the also showed an increase in modulation of riPH (mode = 0.89, 95% HDI = [0.17, 1.49], oHR = 0%; figure 4.7C).

Mediation analysis for the intensity of riPE Sensitivity to changes in SoA modulated riPH (mode = 0.52, 95% HDI = [0.11, 0.97], oHR = 0%, figure 4.7D). Sensitivity to riPE was also shown to increase the intensity of riPE (mode = 0.62, 95% HDI = [0.33, 0.89], oHR = 0%; figure 4.7E), and in a similar fashion being sensitive to both changes in SoA and riPE, potentially had a compounding effect on riPE (mode = 1.43, 95% HDI = [0.70, 2.12], oHR = 0%; figure 4.7F).

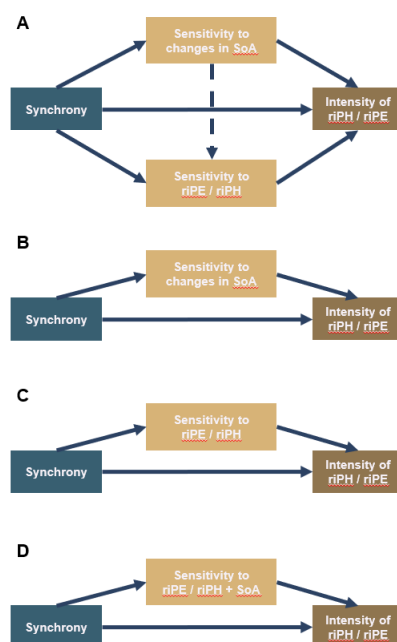


Figure 4.6 Mediation analysis for the effect of sensitivity to different sensations

(A) Hypothesized relationship between synchrony, riPH, riPE, and changes in SoA. This model could not be assessed here as a temporal structure would be necessary to disentangle the second-level of the mediation analysis. **(B)** Mediation analysis used to estimate the posterior for the mediator of changes in SoA for riPH or riPE – also used to establish the prior for SoA for the final model. This was performed in a small subset of 50 participants. **(C)** Mediation analysis used to estimate the posterior for the mediator for sensitivity to riPH or riPE for riPE or riPH, respectively – also used to establish the prior for riPH/riPE for the final model. This was performed in a small subset of participants. **(D)** Final mediation analysis which included 4 levels (not being sensitive to any sensation, being sensitive to riPE/riPH, changes in SoA, or both), for riPH or riPE, respectively.

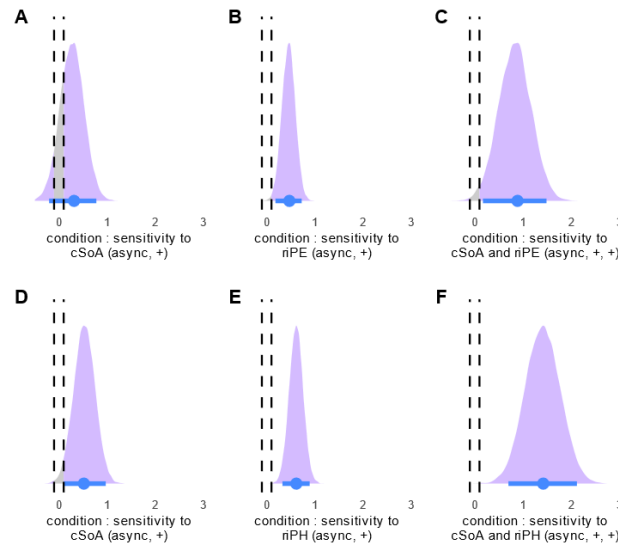


Figure 4.7 Posterior estimates for the sensitivity mediators

For all posteriors estimates the mode is represented by the blue circle, and the 95% HDI by the associated blue line. Distributional information is in purple, and turns to grey if overlapping with the ROPE (within dashed lines). Mediators for riPH: (A) sensitivity to changes in SoA, (B) sensitivity to riPE, (C) both. For riPE: (D) sensitivity to changes in SoA, (E) sensitivity to riPH, (F) both.

4.1.4 Discussion

With this mega-analysis of 26 experiments counting 580 participants, we aimed at investigating multiple aspects of a paradigm which can successfully induce a clinically relevant hallucination, PH, under controlled laboratorial conditions, hence circumventing many limitations in the study of hallucinations. We were able to accurately estimate a medium effect for the main experimental factor used to induce PH: asynchronous robotic sensorimotor stimulation. Importantly, we inferred that some experimental parameters (body location, robot-type) changed the general intensity of riPH, while others (previous exposure, cognitive load) might have a nefarious effects for PH induction. We further determined that schizotypal traits measured as delusional ideations are an important factor for the sensitivity to riPH. Finally we identified the underlying distribution of riPH ratings and we propose a hierarchical relationship between changes in SoA, riPE, and riPH.

4.1.4.1 Relevance of temporal mechanisms for the induction of PH

The temporal feature of the robotic sensorimotor stimulation has consistently been shown as critical for riPH (Bernasconi et al., 2021; Blanke et al., 2014; Orepic et al., 2021; Salomon et al., 2020; Serino et al., 2021). Our analysis confirms that the introduction of an asynchronous sensorimotor stimulation leads to an increase in the intensity of riPH, as compared to the synchronous one, with an effect of medium magnitude, confirming that it's not only tactile stimulation that induces PH. Previous work has also proposed it is indeed the temporal conflict itself that leads to self-related sensorimotor processing errors which in turn lead to riPH (Bernasconi et al., 2021; Blanke et al., 2014). This process of erroneous self-related processing

is hypothesized to be analogous to what is believed to occur naturally in patients with PH (Arzy et al., 2006; Blanke et al., 2003; Brugger et al., 1997). Bernasconi and colleagues (2021), presented compelling evidence for this hypothesis and for the setup being able to mimic the actual symptomatic hallucination, by identifying a group of Parkinson patients suffering from symptomatic PH, which were more sensitive to riPH than patients without symptomatic PH, during stimulation in the asynchronous condition only.

4.1.4.2 Influence of experimental factors on riPH

Given the variability in the induction of PH across different studies, here we aimed at estimating the effect of experimental parameters, other than asynchrony, that varied across experiments. Two factors, body location and robot-type, changed the overall intensity of riPH but did not impact the effect of temporal conflict (i.e. still producing a riPH).

In recent studies, researchers investigated how shifting the sensorimotor conflicts from the trunk to the periphery would impact riPH ratings (Faivre, et al. 2020, Bernasconi, et al, in prep., Franza, et al. in prep.). Here we have estimated that this effect (body location) produces a sizeable decrease in the ratings of riPH, even when controlling for riPH in the front rather than behind the participant. However, despite this decrease in riPH intensity, the effect of asynchrony is preserved given that there's no interaction between body location and asynchrony. Hence, shifting sensorimotor conflicts to the hand, induces a more residual riPH in the asynchronous condition, than the one achieved through trunk sensorimotor conflicts. Further studies in clinical populations with symptomatic PH are needed to understand if this residual induction is still comparable to the PH experienced in daily life.

Another parameter affecting riPH was robot-type, which showed that stronger effect were produced with the classical standing robot compared to the MR-compatible robot. As for body location, no interaction between the predictors of asynchrony and robot-type was observed, meaning that despite a decrease in riPH intensity when using the MRI-robot, the effect of asynchrony was preserved. Given that little can be done to improve MRI settings, researchers should consider an increase in their population sample for MRI experiment, in order to achieve adequate statistical power.

We also highlight two other factors, cognitive load and previous exposure, which carry significant importance for riPH in experiments assessing other aspects of cognition in parallel. The current analysis did not find evidence supporting that any of these parameters impacted riPH, suggesting that tasks can be developed in parallel with or before PH-induction without compromise. However, we do note a consideration regarding both parameters. Given that we did not include in this analysis factors reflecting the intensity or duration of cognitive load or previous exposure, respectively, it could be that an interaction exists but is itself modulated, for example, by the intensity of the cognitive load. Similarly for previous exposure, it could be that longer and more strenuous periods of robot manipulation prior to the assessments hinder riPH (we anecdotally report that experiments with longer periods of previous exposure – e.g. Orepic et al. 2021, experiment 2 and Stripeikyte et al. pilot data: 1 hour – did have smaller effect sizes than other with smaller periods - e.g. Serino et al. 2021: 20 minutes). Experimenters should hence proceed with caution, despite the results of this mega-analysis favoring the null hypothesis.

4.1.4.3 The impact of demographic and trait characteristics on riPH

We have also explored how age, handedness and delusional ideations, modulate riPH ratings.

We have found that PDI, a measure of participants' schizotypal traits and symptoms, was the only parameter to have a clear positive modulation of riPH ratings with increasing PDI scores. While no interaction was observed with asynchrony, suggesting intensity is not affected by delusional ideation, when breaking down this effect for the different values of the Likert's scale, we observed that the extremes of the scale are more sensitive to changes in PDI score. Especially, the lowest value (0: "not at all"), had high likelihood for participants with low PDI scores and sharply decreased with higher PDI scores, while the highest value (6: "completely agree"), showed an opposite pattern of decreased likelihood in participants with low PDI scores, followed by a steep increase with higher PDI scores. Participants with higher delusional ideation scores, might hence be more prone to be generally sensitive to PH-induction, than those with lower scores. Other studies using similar sensorimotor conflicts, albeit visuo-tactile, and to rather induce an illusion, the Rubber Hand Illusion (Botvinick and Cohen, 1998), have also shown a significant impact of delusion proneness on the induction ratings of illusory bodily states. The susceptibility to the RHI has been linked to positive (but not to negative) psychotic traits (Germine et al., 2013), with later work linking positively the perceived intensity of the RHI, to paranoid ideations and psychoticism in healthy individuals (Kállai et al., 2015). Recent work in a large sample of healthy individuals, also assessed with the PDI, reached similar conclusions with delusional proneness being positively linked to RHI intensity (Louzolo et al., 2015).

While age did not produce any meaningful change in riPH intensity, a decomposition of this effect for the different values of the Likert's scale, revealed a sharp increase in the likelihood of rating the lowest values (0: "not at all"), with increased age, whereas higher values decreased their likelihood with age. It is possible, that younger people have a more uniform distribution in terms of sensitivity to riPH, and older individuals are both less sensitive to it, and more likely to report lower intensities as compared to younger individuals. This would be in line with recent studies with the RHI, which showed decreased induction performance with age. Specifically, more vivid inductions, higher percentages of respondents, and even earlier illusion onsets have been confirmed for younger individuals (Ferracci and Brancucci, 2019; Kállai et al., 2017). This could be an important factor to take into account, as the clinical populations suffering from sPH and studied for riPH are generally older (Bernasconi et al., 2021; Salomon et al., 2020), than the young healthy populations investigated in other riPH studies (Dhanis et al., 2022; Faivre et al., 2020; Orepic et al., 2021; Serino et al., 2021).

4.1.4.4 Population-level bias in interpreting the riPH rating scale

A secondary objective of this mega-analysis was to estimate the distribution underlying the ordinal scale in which riPH is rated. While this might seem like a purely technical endeavor to some, this will be consequential for future studies investigating riPH, as knowing the correct parameters for the underlying distribution can boost the power of the analysis.

4.1.4.5 Robot-induced PH and accompanying sensations

Finally, we aimed at elaborating on a relationship between riPH, riPE, and changes in SoA, hypothesizing that specifically changes in SoA, would be at the base of the two previous sensations. To put such analysis in perspective it is important to reconsider three points. One, PH is a hallucination that is considered to stem from a distortion of own-body perception (Blanke et al., 2008). Two, PE on the other hand is by definition the sensation that someone else is the agent of your own actions (Mlakar et al., 1994), and involves a process of alienation of said actions. And finally three, SoA, assessed here as Loss of Agency, is conceptually close to PE, but does not require the alienation process, as it simply refers to the process of no longer considering yourself the agent of your actions (Haggard and Chambon, 2012), without necessarily attributing them to someone else. A particular focus is put on these as given their definitions and our results, we hypothesize that in the current robotic setup and for PH induction in healthy individuals, there might be an underlying hierarchical relationship between these three phenomena.

The mediation analysis evidenced that the sensorimotor conflicts used can induce changes in SoA, in particular with the asynchronous condition resulting in loss of agency. Many other studies have confirmed the effect of asynchronous sensorimotor conflicts in perceived changes in SoA (for a meta-analysis on motor control and agency: (Zito et al., 2020)). We further propose that it might be the case that the experience of riPE depends on loss of agency. This would be a logical conclusion from their definitions (PE explicitly requires an alienation of own's one actions after losing ownership for them, i.e. losing agency), further supported by our data, but would also be in line with clinical observations, given that PE has already been hypothesized to have a dependency on agency deficits in schizophrenic patients (Moore, 2016). The dependency of riPH on a negative change in SoA was however not clear from our data, and while it could be that there is some effect of losing agency in the induction of riPH, here our estimates could not conclusively arrive at a supporting or countering decision.

Interestingly, the sensitivity to riPE did make participants experience a more intense riPH (and vice-versa). This could point to a synergetic relationship between riPH and riPE. Neuroimaging data has in fact indicated that while two distinct processes of changes in transition probabilities between brain networks characterize sensitivity to riPH and riPE (Dhanis et al., 2022), at least the riPE process is also observed for riPH (albeit not in a significant manner to explain riPH).

Finally, being sensitive to both changes in SoA and riPE/riPH, seemed to impact riPH/riPE in a compounding manner when compared to any of the former two isolated. This could either support our initial hypothesis of a two-step mediation process between SoA, riPE/riPH, and intensity of riPH/riPE (figure 6A), or it could support a parallel mediation process. Albeit not compared directly in this analysis our results do seem to suggest a higher compounding effect for riPE, than riPH. Previous work has highlighted dependencies of passivity experiences on alienation processes (ref) , which could explain this.

This analysis includes some particular limitations. Our initial hypothesis (figure 6A) is not directly testable as a two-step mediation procedure requires a temporal structure to disambiguate the order of mediation. This would never be possible in this scenario. Hence, we

tried to create a predictor which to an extent encodes, but does not enforce such a relationship. Another limitation was that the amount of data for changes in SoA, limited the power of this analysis, and forced us to establish separate priors to improve the quality of the last model. While we guaranteed that this was performed in separate batches of participants representing all experiments, we naturally cannot guarantee that changes in SoA did not occur, in the studies where it was not assessed.

4.1.4.6 Conclusion

In brief, with this mega-analysis we have accurately estimated how various experimental parameters modulate riPH in healthy individuals, advised on the use of such parameters, and identified a subset of the population with higher delusional ideations which might be more sensitive to riPH. We based these contributions on strong statistical tools, which we hope will be useful for future studies assessing robot-induced PH. All these results, although in healthy individuals, are highly informative, for a setup that is now being extended to various clinical applications, including Parkinson's (Bernasconi et al., 2021). Finally we have also extended the results of this mega-analysis and other studies (Dhanis et al., 2022) to propose a hierarchical relationship between changes in SoA and the sensitivity to both riPE and riPH, as well as a synergetic relationship between riPE and riPH.

4.1.5 Methods

4.1.5.1 Studies and data

The studies included here were almost all conducted in the Laboratory of Cognitive Neurosciences and are either published, in submission process, or constitute pilot data that were not published, and which we released with the current analysis in an effort to counter publication bias. For the published studies, they can be retrieved from Pubmed, and Web of Science, using the query arguments: (presence) AND (hallucination OR illusion OR feeling) AND (robot OR robotically) AND (sensorimotor) AND (conflict). Following the PICO framework, studies were included when the following criteria were met:

- Population: Healthy individuals
- Intervention: Robotically controlled induction of sensorimotor conflicts
- Comparison condition: Exposure to different experimental conditions
- Outcome: Strength of induced illusory perceptions, including PH, PE, changes in SoA, amongst others (as assessed through follow-up questionnaires using 7-point Likert scales).

For each study, we retrieved the raw questionnaire ratings for every participant, the experimental parameters of that experimental paradigm (described further below), and participants' demographics (described further below).

4.1.5.2 General experimental setup

All the experiments included in this mega-analysis used the same base for the experimental setup. Participants were blindfolded with an eye-mask and white noise was presented through headphones to isolate them from surrounding noise. While isolated, participants had a handle attached to their right index finger and moved the front part of a robotic system (front-robot) by performing poking movements. At the same time, these movements were transmitted and reproduced by a back-robot, onto the participant's backs. To assess the subjective experience induced by the sensorimotor robotic stimulation, each participant was asked to respond to a questionnaire immediately after each experimental manipulation. For each questionnaire item, participants had to rate on a 7-point Likert's scale (from 0 being "not at all", to 6, being "very strongly"), the intensity of the subjective experience. The following two items were included in every experiment: "I felt as if someone was behind me", to assess the illusory sensation of riPH, and "I felt as if someone else was touching my body", to assess the induction of PE. Aside from these, questions assessing other subjective sensations (i.e. LoA, the experience of self-touch, anxiety), and different control questions were used across the different studies.

4.1.5.3 Experimental parameters

Since the original PH induction study (Blanke et al., 2014), the various follow-up studies have used at times different experimental parameters to either, adapt the paradigm to the MRI, explore different control conditions, or to investigate cognitive functions being explored concomitantly with PH induction. Here, we describe all parameters that vary within and between experiments.

Within experiment: Every experiment repeated the period of robot manipulation at least twice, having one period of robot manipulation with the front- and back-robot working in the synchronous condition (where participants receive the feedback on their back in synchrony with their movements), and another period where the robots were in an asynchronous condition, introducing a delay of 500 ms between the movements in the front and the feedback received on the back. The order of the two experimental conditions was always randomized between participants. To investigate the possible effect on the questionnaires ratings of the order two experimental conditions we investigated (Order) in the current mega-analysis. Some experiments also included two more repetition periods, where the feedback was given to the hand rather than the back, as a control condition (Body part: back or hand). This is because the presentation of the tactile feedback specifically to the back, has been hypothesized to be important for PH induction, with some researchers exploring if changing the body location of the tactile feedback to a more peripheral location would also result in PH induction (Faivre et al., 2020).

Between experiments: Due to differences in experimental setup specific to each study, several parameters varied across the experiments: the duration of the robot manipulation period (duration), the type of robot used (robot-type, standing robot or MRI robot), previous exposure to the robotic stimulation (i.e., some experiments had participants use the robot for another task, before PH induction), the use of a virtual back in front of the participant (i.e. front-robot resisted the participant's movement in front as if they were touching someone's back in front of them), and the presence of concomitant tasks during the induction of PH (cognitive load). In this analysis the factor cognitive load is encoded as a binary value (Yes/No). This simplification implies that we considered any concomitant task as cognitive load, although the actual cognitive load might have varied between experiments.

4.1.5.4 Participants' demographics and trait characteristics

In this mega-analysis, we have also investigated if participants' characteristics influenced the induction of rIPH and other subjective experiences. This included their age, handedness score as assessed by the Edinburgh Handedness Inventory (Oldfield, 1971), and delusional ideation score as assessed by the Peter's Delusional Inventory (Peters et al., 2004). Sex at birth was included as a covariate of no-interest.

4.1.5.5 Statistical analysis

We conducted a Bayesian Estimation analysis to assess the effect of asynchrony and different experimental parameters, on the ratings of riPH and accompanying illusions, following literature recommendations (John K Kruschke and Liddell, 2018b, 2018a; Liddell and Kruschke, 2018). All analyses were performed with the use of the brms package (Bürkner, 2017) available for R (version 4.0.5).

To estimate the effect of the experimental parameters and general demographic data (sex at birth and age) on the induced illusions we first fitted a full model on the data (FM1), as seen below:

$$\begin{aligned} \text{Rating} \sim & \text{Asynchrony} + \text{BodyLocation} + \text{RobotType} + \text{Previous Exposure} + \text{Order} + \text{Duration} + \text{Cognitive Load} \\ & + \text{Virtual Back} + \text{Synchrony: Body Location} + \text{Synchrony: RobotType} + \text{Synchrony: Cognitive Load} \\ & + \text{Synchrony: Previous Exposure} + \text{Synchrony: Order} + \text{Synchrony: Duration} + \text{Synchrony: VirtualBack} \\ & + \text{Age} + \text{Age: Asynchrony} + \text{Sex at birth} + (1 + \text{Synchrony} | \text{Experiment}) \end{aligned}$$

The inclusion of the interactions between synchrony and body part, cognitive load, previous exposure, and order respectively, relates to our experimental question that those parameters might influence the effect of synchrony, which had been in the original and follow-up studies the main parameter enhancing riPH. A random intercept term for experiment was also included, to capture any potential variability that is not expressed in the model. A random slope for Synchrony varying with experiment was included, to explore the variation of the effect of Synchrony across the many experiments. Regarding model estimation, our model did not include informed priors (flat priors were used), and a cumulative probit link-function with a flexible threshold was used.

Two more models were fitted separately in smaller populations, to account for the effects of EHI and PDI scores, respectively. The models were the same as FM1, but included added term for EHI and the interaction between asynchrony and EHI (FM2), or the same but applied to PDI (FM3). These had to be assessed separately given that EHI and PDI were not necessarily available for the same smaller sub-population (albeit some overlap).

Once the beta estimates and high-density intervals (HDI) for the effect of the different predictors were estimated, we used the Region of Practical Equivalence (ROPE) procedure to establish which experimental parameters modulated the ratings of each illusion. In sum, this statistical procedure establishes a region around the null value, in this case 0, which is considered to be practically equivalent to the null value (John K Kruschke and Liddell, 2018a). We have considered this ROPE to extend ± 0.1 standard deviations (of the global distribution of the ratings for each assessed illusion) from the null value, which is equivalent to considering very small effect sizes as no practical effect. If the HDI of an estimate falls completely within this ROPE, then we considered this parameter to be practically equivalent to the null value, and hence that it does not modulate the ratings of the illusions. Otherwise, for a specific parameter, if its HDI fell completely outside the ROPE, we considered that this parameter has an effect different than the null value, and hence modulates the experience of the induced illusion. In the case the estimated beta of a parameter falls outside the ROPE, but the HDI of this estimate overlaps with the ROPE, we remain undecided (i.e. cannot come to a definite conclusion) on whether that parameter modulates the experience of the induced illusion.

differently than the null value. We provided a measurement of overlap between the HDI and ROPE, here referred to as oHR.

4.1.5.6 Estimation of the underlying PH rating scale distribution

The estimation of the distribution that underlies the 7-point Likert ordinal scale used to rate riPH, comes as a direct consequence of the type of model used above (cumulative probit link-function with a flexible threshold). We choose to highlight this methodological aspect here, given its importance for future studies.

The model used, included 6 parameters that estimated the thresholds between each point of the rating scale (i.e. between “2” and “3”). These parameters identified the underlying distribution. For the purposes of more adequate visualization, we then constructed a generative model, that used only the these 6 intercept parameters. Such a model outputs the data, in a way that is as close to the participants’ internal biases towards the “scale + assessment” as it can be (population-level). For comparison purposes, one can think to then compare this underlying distribution, to an hypothesized metric distribution where all values of the scale are equally spaced.

4.1.5.7 Mediation analysis and the relationship between sensations’ induction

To assess potential relationships between the induction of riPH and the typically accompanying sensations of riPE and changes in SoA, we modelled how sensitivity to changes in SoA, and to riPH or riPE, affected the intensity of riPE or riPH, respectively (figure 6). While sensitivity to induction refers to the induction of these illusions regardless of their intensity (e.g. any positive score difference between asynchronous and synchronous conditions is coded as being sensitive), intensity of induction refers to the actual difference between the scores in the asynchronous versus synchronous conditions. We did so using an approach akin to mediation analysis, with special considerations. A typical mediation analysis would assess if for example asynchrony modulates riPH, and if it further modulates changes in SoA potentiating further modulation of riPH. Here we were additionally interested in knowing if for example sensitivity to riPE would also modulate riPH intensity. Knowing that this variable is not independent from sensitivity to changes in SoA added an additional problem (discussed here: VanderWeele and Vansteelandt 2012). To circumvent this problem we modelled a three-way interaction between asynchrony, sensitivity to changes in SoA, and sensitivity to riPH or riPE (depending on if the dependent variable was intensity of riPE or riPH). This approach provides a better clarification of the possible interactions between these variables, however, includes a directionally uninteresting effect which is the modulation of changes in SoA due to riPE/riPH sensitivity. Appropriate circumvention of this problem would have to be performed through Structural Equation modelling (as done here: Votruba-Drzal, et al 2014), however this was outside the scope of the current paper.

4.1.6 References

- Allefeld C, Pütz P, Kastner K, Wackermann J. 2011. Flicker-light induced visual phenomena: Frequency dependence and specificity of whole percepts and percept features. *Conscious Cogn* 20:1344–1362. doi:10.1016/j.concog.2010.10.026
- Arzy S, Seeck M, Ortigue S, Spinelli L, Blanke O. 2006. Induction of an illusory shadow person. *Nature* 443:287. doi:10.1038/443287a
- Baggott MJ, Siegrist JD, Galloway GP, Robertson LC, Coyle JR, Mendelson JE. 2010. Investigating the mechanisms of hallucinogen-induced visions using 3,4-methylenedioxymphetamine (MDA): A randomized controlled trial in humans. *PLoS One* 5. doi:10.1371/journal.pone.0014074
- Bernasconi F, Blondiaux E, Potheegadoo J, Stripeikyte G, Pagonabarraga J, Bejr-kasem H, Bassolino M, Akselrod M, Martinez-horta S, Sampedro F, Hara M, Horvath J, Franza M, Konik S, Bereau M, Ghika J, Burkhard PR, Ville D van de, Faivre N, Rognini G, Krack P. 2021. Robot-induced hallucinations in Parkinson ' s disease depend on altered sensorimotor processing in fronto-temporal network. *Sci Transl Med* 8362:1–13.
- Berrios GE, Marková IS. 2002. Conceptual issues. *Biol Psychiatry* 11–32.
- Bilder RM, Goldman RS, Robinson D, Reiter G, Bell L, Bates JA, Pappadopulos E, Willson DF, Maria J, Alvir J, Woerner MG, Geisler S, Kane JM, Lieberman JA. 2000. Neuropsychology of First-Episode Schizophrenia: Initial Characterization and Clinical Correlates, *Am J Psychiatry*.
- Blanke O, Arzy S, Landis T. 2008. Chapter 22 Illusory reduplications of the human body and self. *Handb Clin Neurol* 88:429–458. doi:10.1016/S0072-9752(07)88022-5
- Blanke O, Ortigue S, Coeytaux A, Martory MD, Landis T. 2003. Hearing of a presence. *Neurocase* 9:329–339. doi:10.1076/neur.9.4.329.15552
- Blanke O, Pozeg P, Hara M, Heydrich L, Serino A, Yamamoto A, Higuchi T, Salomon R, Seeck M, Landis T, Arzy S, Herbelin B, Bleuler H, Rognini G. 2014a. Neurological and robot-controlled induction of an apparition. *Current Biology* 24:2681–2686. doi:10.1016/j.cub.2014.09.049
- Blanke O, Pozeg P, Hara M, Heydrich L, Serino A, Yamamoto A, Higuchi T, Salomon R, Seeck M, Landis T, Arzy S, Herbelin B, Bleuler H, Rognini G. 2014b. Neurological and Robot-Controlled Induction of an Apparition. *Current Biology* 24:2681–2686. doi:10.1016/j.cub.2014.09.049
- Botvinick M, Cohen J. 1998. Rubber hands feel touch that eyes see. *Nature Scientific Correspondance* 391:756.
- Brugger P, Regard M, Landis T. 1997. Illusory reduplication of one's own body: Phenomenology and classification of autoscopic phenomena. *Cogn Neuropsychiatry* 2:19–38. doi:10.1080/135468097396397
- Brugger P, Regard M, Landis T. 1996. Unilaterally felt "presences": The neuropsychiatry of one's invisible doppelgänger. *Neuropsychiatry Neuropsychol Behav Neurol* 9:114–122.

Part III: Quantifying and optimizing PH induction

- Brugger P, Regard M, Landis T, Oelz O. 1999. Hallucinatory experiences in extreme-altitude climbers. *Neuropsychiatry Neuropsychol Behav Neurol* 12:67–71.
- Bürkner PC. 2017. brms: An R package for Bayesian multilevel models using Stan. *J Stat Softw* 80. doi:10.18637/jss.v080.i01
- Case LK, Solcá M, Blanke O, Faivre N. 2019. Disorders of body representation, Multisensory Perception: From Laboratory to Clinic. Elsevier Inc. doi:10.1016/B978-0-12-812492-5.00018-8
- Chan D, Rossor MN. 2002. “-but who is that on the other side of you?” Extracampine hallucinations revisited. *Lancet* 360:2064–2066. doi:10.1016/S0140-6736(02)11998-2
- de Haan EH, Nys GM, van Zandvoort MJ, Ramsey NF. 2007. The physiological basis of visual hallucinations after damage to the primary visual cortex.
- Dhanis H, Blondiaux E, Bolton T, Faivre N, Rognini G, van de Ville D, Blanke O. 2022. Robotically-induced hallucination triggers subtle changes in brain network transitions. *Neuroimage* 248:118862. doi:10.1016/j.neuroimage.2021.118862
- Dickersin K. 2005. Publication Bias: Recognizing the Problem, Understanding Its Origins and Scope, and Preventing Harm Publication Bias in Meta-Analysis. John Wiley & Sons, Ltd. pp. 9–33. doi:https://doi.org/10.1002/0470870168.ch2
- Ehrsson HH, Holmes NP, Passingham RE. 2005. Touching a rubber hand feeling of body ownership is associated with activity in multisensory brain areas. *The Journal of Neuroscience* 25:10564–10573. doi:10.1523/JNEUROSCI.0800-05.2005
- Faivre N, Vuillaume L, Bernasconi F, Salomon R, Blanke O, Cleeremans A. 2020. Sensorimotor conflicts alter metacognitive and action monitoring. *Cortex* 124:224–234. doi:10.1016/j.cortex.2019.12.001
- Fénelon G, Soulas T, de Langavant LC, Trinkler I, Bachoud-Lévi AC. 2011. Feeling of presence in Parkinson’s disease. *J Neurol Neurosurg Psychiatry* 82:1219–1224. doi:10.1136/jnnp.2010.234799
- Ferracci S, Brancucci A. 2019. The influence of age on the rubber hand illusion. *Conscious Cogn* 73. doi:10.1016/j.concog.2019.05.004
- Ffytche H D, Creese B, Politis M, Chaudhuri KR, Weintraub D, Ballard C, Aarsland D. 2017. The psychosis spectrum in Parkinson disease. *Nat Rev Neurol*. doi:10.1038/nrneurol.2016.200
- Flint A, Loh J, Brust J. 2005. Vivid visual hallucinations from occipital lobe infarction.
- Geiger J. 2009. The third man factor. Hachette Books.
- Germine L, Benson TL, Cohen F, Hooker CIL. 2013. Psychosis-proneness and the rubber hand illusion of body ownership. *Psychiatry Res* 207:45–52. doi:10.1016/j.psychres.2012.11.022
- Haggard P, Chambon V. 2012. Sense of agency. *Current Biology* 22:R390–R392. doi:10.1016/j.cub.2012.02.040
- Harris R, Gurel L. 2012. A study of ayahuasca use in North America. *J Psychoactive Drugs* 44:209–215. doi:10.1080/02791072.2012.703100

Part III: Quantifying and optimizing PH induction

- Kállai J, Hegedüs G, Feldmann Á, Rózsa S, Darnai G, Herold R, Dorn K, Kincses P, Csathó Á, Szolcsányi T. 2015. Temperament and psychopathological syndromes specific susceptibility for rubber hand illusion. *Psychiatry Res* 229:410–419. doi:10.1016/j.psychres.2015.05.109
- Kállai J, Kincses P, Lábadi B, Dorn K, Szolcsányi T, Darnai G, Hupuczi E, Janszky J, Csathó Á. 2017. Multisensory integration and age-dependent sensitivity to body representation modification induced by the rubber hand illusion. *Cogn Process* 18:349–357. doi:10.1007/s10339-017-0827-4
- Kruschke John K., Liddell TM. 2018. The Bayesian New Statistics: Hypothesis testing, estimation, meta-analysis, and power analysis from a Bayesian perspective. *Psychon Bull Rev* 25:178–206. doi:10.3758/s13423-016-1221-4
- Kruschke John K, Liddell TM. 2018a. The Bayesian New Statistics : Hypothesis testing , estimation , meta-analysis , and power analysis from a Bayesian perspective 178–206. doi:10.3758/s13423-016-1221-4
- Kruschke John K, Liddell TM. 2018b. Bayesian data analysis for newcomers 155–177. doi:10.3758/s13423-017-1272-1
- Kühn S, Gallinat J. 2012. Quantitative meta-analysis on state and trait aspects of auditory verbal hallucinations in schizophrenia. *Schizophr Bull* 38:779–786. doi:10.1093/schbul/sbq152
- Lenggenhager B, Taid T, Metzinger T, Blanke O. 2007. Video Ergo Sum: Manipulating Bodily Self-consciousness. *Science* (1979) 317:1093–1096. doi:DOI: 10.1126/science.1143439
- Lenka A, Pagonabarraga J, Pal PK, Bejr-kasem H, Kulisvesky J. 2019. Minor hallucinations in Parkinson disease A subtle symptom with major clinical implications 1–9. doi:10.1212/WNL.00000000000007913
- Liddell TM, Kruschke JK. 2018. Analyzing ordinal data with metric models: What could possibly go wrong? *J Exp Soc Psychol* 79:328–348. doi:10.1016/j.jesp.2018.08.009
- Llorca PM, Pereira B, Jardri R, Chereau-Boudet I, Brousse G, Misdrahi D, Fénelon G, Tronche AM, Schwan R, Lançon C, Marques A, Ulla M, Derost P, Debilly B, Durif F, de Chazeron I. 2016. Hallucinations in schizophrenia and Parkinson’s disease: An analysis of sensory modalities involved and the repercussion on patients. *Sci Rep* 6. doi:10.1038/srep38152
- Louzolo A, Kalckert A, Petrovic P. 2015. When passive feels active - Delusion-proneness alters self-recognition in the moving rubber hand illusion. *PLoS One* 10:1–12. doi:10.1371/journal.pone.0128549
- Messner R. 2016. *Naked Mountain: Nanga Parbat, Brother, Death, Solitude*. Crowood.
- Mlakar J, Jensterle J, Frith CD. 1994. Central monitoring deficiency. *Psychol Med* 24:557–564.
- Moher D, Liberati A, Tetzlaff J, Altman DG, Group TP. 2009. Preferred Reporting Items for Systematic Reviews and Meta-Analyses : The PRISMA Statement 6. doi:10.1371/journal.pmed.1000097
- Moore JW. 2016. What is the sense of agency and why does it matter? *Front Psychol* 7:1–9. doi:10.3389/fpsyg.2016.01272
- Myhren H, Ekeberg Ø, Tøien K, Karlsson S, Stokland O. 2010. Posttraumatic stress, anxiety and depression symptoms in patients during the first year post intensive care unit discharge.

Part III: Quantifying and optimizing PH induction

- Nagahama Y, Okina T, Suzuki N, Matsuda M. 2010. Neural correlates of psychotic symptoms in dementia with Lewy bodies. *Brain* 133:557–567. doi:10.1093/brain/awp295
- Nagahama Y, Okina T, Suzuki N, Matsuda M, Fukao K, Murai T. 2007. Classification of Psychotic Symptoms in Dementia With Lewy Bodies.
- Nisbett RE, Wilson TD. 1977. Telling more than we can know: Verbal reports on mental processes. *Psychol Rev* 84:231–259. doi:10.1037/0033-295X.84.3.231
- Oldfield RC. 1971. The assessment and analysis of handedness: The Edinburgh Inventory. *Neuropsychologia* 9:97–113. doi:10.1007/978-0-387-79948-3_6053
- Orepic P, Rognini G, Kannape OA, Faivre N, Blanke O. 2021. Sensorimotor conflicts induce somatic passivity and louden quiet voices in healthy listeners. *Schizophr Res* 231:170–177. doi:10.1016/j.schres.2021.03.014
- Pagonabarraga J, Martinez-Horta S, Fernández de Bobadilla R, Pérez J, Ribosa-Nogué R, Marín J, Pascual-Sedano B, García C, Gironell A, Kulisevsky J. 2016. Minor hallucinations occur in drug-naïve Parkinson's disease patients, even from the premotor phase. *Movement Disorders* 31:45–52. doi:10.1002/mds.26432
- Pagonabarraga J, Soriano-Mas C, Llebaria G, López-Solà M, Pujol J, Kulisevsky J. 2014. Neural correlates of minor hallucinations in non-demented patients with Parkinson's disease. *Parkinsonism Relat Disord* 20:290–296. doi:10.1016/j.parkreldis.2013.11.017
- Peters E, Joseph S, Day S, Garety P. 2004. Measuring delusional ideation: The 21-item Peters et al. Delusions Inventory (PDI). *Schizophr Bull* 30:1005–1022. doi:10.1093/oxfordjournals.schbul.a007116
- Picard F. 2010. Epileptic feeling of multiple presences in the frontal space. *Cortex* 46:1037–1042. doi:10.1016/j.cortex.2010.02.002
- Potheegadoo J, Dhanis H, Horvath J, Burkhard PR, Blanke O. 2022. Presence Hallucinations during Locomotion in Patients with Parkinson's Disease. *Mov Disord Clin Pract*. doi:10.1002/mdc3.13367
- Ravina B, Marder K, Fernandez HH, Friedman JH, McDonald W, Murphy D, Aarsland D, Babcock D, Cummings J, Endicott J, Factor S, Galpern W, Lees A, Marsh L, Stacy M, Gwinn-Hardy K, Voon V, Goetz C. 2007. Diagnostic criteria for psychosis in Parkinson's disease: Report of an NINDS, NIMH Work Group. *Movement Disorders* 22:1061–1068. doi:10.1002/mds.21382
- Salomon R, Progin P, Griffa A, Rognini G, Do KQ, Conus P, Marchesotti S, Bernasconi F, Hagmann P, Serino A, Blanke O. 2020. Sensorimotor induction of auditory misattribution in early psychosis. *Schizophr Bull* 46:947–954. doi:10.1093/schbul/sbz136
- Serino A, Pozeg P, Bernasconi F, Solcà M, Hara M, Progin P, Stripeikyte G, Dhanis H, Salomon R, Bleuler H, Rognini G, Blanke O. 2021. Thought consciousness and source monitoring depend on robotically controlled sensorimotor conflicts and illusory states. *iScience* 24. doi:10.1016/j.isci.2020.101955
- Sharpe d. 1997. Of apples and oranges, file drawers and garbage: why validity issues in meta-analysis will not go away, *Clinical psychology review*.

Part III: Quantifying and optimizing PH induction

Timmermann C, Roseman L, Williams L, Erritzoe D, Martial C, Cassol H, Laureys S, Nutt D, Carhart-Harris R. 2018. DMT models the near-death experience. *Front Psychol* 9. doi:10.3389/fpsyg.2018.01424

Wackermann J, Pütz P, Allefeld C. 2008a. Ganzfeld-induced hallucinatory experience, its phenomenology and cerebral electrophysiology. *Cortex* 44:1364–1378. doi:10.1016/j.cortex.2007.05.003

Wackermann J, Pütz P, Allefeld C. 2008b. Ganzfeld-induced hallucinatory experience, its phenomenology and cerebral electrophysiology. *Cortex* 44:1364–1378. doi:10.1016/j.cortex.2007.05.003

Whitfield-Gabrieli S, Ford JM. 2012. Default mode network activity and connectivity in psychopathology. *Annu Rev Clin Psychol* 8. doi:10.1146/annurev-clinpsy-032511-143049

Wood RA, Hopkins SA, Moodley KK, Chan D, Gonzalez-billault C. 2015. Fifty Percent Prevalence of extracampine hallucinations in Parkinson ' s Disease Patients 6:1–9. doi:10.3389/fneur.2015.00263

Zhang X, Savalei V. 2016. Improving the Factor Structure of Psychological Scales: The Expanded Format as an Alternative to the Likert Scale Format. *Educ Psychol Meas* 76:357–386. doi:10.1177/0013164415596421

Zito GA, Wiest R, Aybek S. 2020. Neural correlates of sense of agency in motor control: A neuroimaging meta-analysis. *PLoS One* 15:1–17. doi:10.1371/journal.pone.0234321

General Discussion

My thesis work revolved around two major points. One, neural correlates of hallucinations in health and disease can be assessed through approaches that resolve the temporal dynamics of brain activity. Two, if the underlying process of a hallucination is known, then it should be controllable, as long as informative feedback is provided to the participants.

I was able to pair these two aspects across different experiments, with a robotic system that is capable of inducing a specific and clinically relevant hallucination: presence hallucination (PH). In doing so, I was able to circumvent a significant problem in hallucination research, which is studying hallucinations directly without the confound of disease trait. Furthermore, using methods that decompose the temporal structure of brain activity, I identified the neural correlates of PH as it was induced, had participants perform neurofeedback to control the neural correlates of PH concomitantly with induction, and studied the neural correlates of hallucinations in PD, directly associating them as well to the induced PH.

In this chapter, I will summarize the findings of my work, relating them with one another as well as broader literature on the topic, and discuss future directions.

5.1 The importance of temporal dynamics

In Study I of this thesis, I re-analyzed neural data from a previous study, which used the MR-compatible robotic system to induce PH in healthy individuals (Bernasconi et al., 2021). One of the findings of this study was that when inducing PH in healthy individuals, certain regions were more active in the asynchronous condition, where PH is induced, versus a control condition. Due to the complex nature of that study and given its focus on brain activity in neurological populations with PH, an unaddressed question was the specific brain activity underlying induction of PH and other accompanying sensations, beyond the brain activity associated to the task. Prominent research in psychosis, was advancing the novel idea at the time, that mechanisms of hallucinations and psychosis were “hiding” in temporal dynamics of brain activity (Menon, 2011; Palaniyappan & Liddle, 2012; Stephan et al., 2009). Moreover, some studies had already found general dysfunctions in interactions within the triple network model (that is, in how the salience network modulates interactions between the DMN and CEN) as well as dysfunctions in salience processing in patients suffering from psychosis (Manoliu

General Discussion

et al., 2014; Moran et al., 2013; Palaniyappan et al., 2013). Hence, with this in mind, I hypothesized whether the neural underpinnings of PH induction could be found in the temporal dynamics of brain activity.

To explore this, I resorted to dynamic functional connectivity (Preti et al., 2017), which can resolve the temporal structure of brain activity, in particular CAPs, which make use of previously known information about brain regions relevant for the task, so called seed regions which are used as a base to study succession of brain networks at relevant timepoints. In our case these seeds were the ones found to be more active in the asynchronous condition in previous work (pSTS and IFG). With this I identified several brain networks active while participants used the PH-inducing robot, in a condition designed to induce PH (asynchronous condition) and a control condition. By studying their interactions in the form of occurrences and transition probabilities, I was able to identify that higher occurrence of one brain network was tied to the asynchronous condition, however this was not yet what was behind riPH, as various sensations are induced in the asynchronous condition (e.g., PE, changes in SoA). Interestingly, underlying the specific sensitivity to riPH, was in fact how all the brain networks changed their transition probabilities to favor a single brain network (the same brain network which occurred more in the asynchronous condition). We called this brain network, the PH-network. Importantly, we made sure to control this for the confound of concomitantly induced sensations, namely passivity experiences (PE). My analysis revealed that in fact while underpinning riPH there was this favoring of transitions to a single brain network, PE was characterized by a different process where the DMN was shunned in the transition space. With this, I also address the problem of “demand characteristics” (i.e. suggestibility), as our participants were blinded to which sensation we were interested in (there were other assessments and control question as well).

Finally, an important aspect for me, given that I was working with an induced hallucination (in healthy individuals in the case of Part I – Study 1), was to put my findings in the perspective of what is known in clinical populations. In other words: *how do the neural correlates of this induced hallucination compare to those of clinical hallucinations?* For this purpose I compare my findings to studies that were able to directly study hallucinations and rid of the confound of disease traits.

Lefebvre and colleagues (2016), who studied periods of hallucinations in schizophrenic patients, identified general mechanisms that were not tied to specific activations or novel brain patterns, but rather to shifts in interaction dynamics across large network in the triple network model. A more recent study in PD identified that “unstable” transitions between states, were underlying visual hallucinations in these patients, whereas such transitions occurred much more rarely in patients without hallucinations (Zarkali et al., 2022). In sum, I consider that the mechanisms I have identified in riPH are very similar in nature to those proposed by theories of network interactions (Menon, 2011), and are in practice similar to those identified in clinical work for clinical hallucinations.

Naturally, what is being argued in here can be extended in full for my NF work (Part I – Study 2), which also identified significant changes in transition dynamics, especially for participants that engaged in successful regulation of the PH-network.

5.1.1 Indirect evidence for hierarchical dynamics: Meta-states

Another implication of my studies is an aspect that goes considerably beyond what I discuss in them, as it would simply fall outside the scope of the specific papers. In Study I, I showed that participants sensitive to riPH exhibited a certain type transition pattern amongst brain states, whereas those sensitive to riPE had another, those that were not to anything had yet another, and so on. I will argue here that this is actually highly supportive of a meta-state organization.

We have discussed brain *states*, in almost all studies presented in this thesis. So what is a *meta-state*? A meta-state refers to hierarchical organization in which two (or potentially more) levels exist. On the bottom there are the observed states, for example brain states of activity as the ones we have been discussing, and higher in the hierarchy are “invisible” meta-states, which do not imply different observed states, but that set the rules between transitions or interactions across states. For example, one study showed that across 820 participants, individuals’ patterns of brain activity during rest followed mostly two types of non-stochastic transitioning patterns (i.e. meta-states; Vidaurre et al., 2017; Figure 5.1). These meta-states were associated with different behavioral characteristics and were heritable. A more mathematical explanation of meta-states can be found in Mucha and colleagues (2010), and further applications to the human neuroimaging in Betzel and Bassett (2017).

What I am arguing here, is that what we observe by studying neural correlates of riPH through brain dynamics and temporal modelling is in fact the nature of these meta-states. That was the aspect that distinguished sensitivity to riPH, riPE, and so on, and thus this strongly suggests that there is evidence for meta-states from that study. I could even add that the dependency observed between PE on PH on Study IV, is further supporting this claim. What, could not have directly been assessed (in Study I) was the relationship between meta-states (the invisible hierarchy). However, note that it was not a goal of Study I to address these further question. We could not have known exactly if the meta-states were circulating between each other, or if a specific meta-state made participants prone to riPH, riPE, or both. If we consider evidence across multiple brain studies, the scales due tip to favor the former

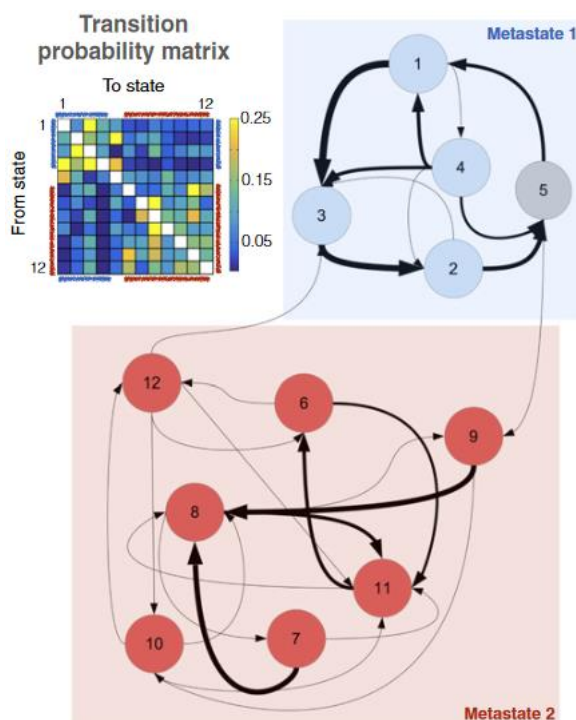


Figure 5.1 Example of meta-states

Meta-states from Vidaurre and colleagues (2017) are shown as an example. A transition probability matrix is shown on the left highlight to groups (i.e. meta-states) with states more likely to transition amongst each other. Those meta-states are shown on the left and bottom in blue and red. While in this example the states which belong to each meta-state are all different, the reader should not assume this is a requirement to establish meta-states. A state can belong to different meta-states.

(Finn et al., 2015; Rosenberg et al., 2016; Sadaghiani et al., 2015; Vidaurre et al., 2017). The same, could be said about the NF experiment (Study II), as participants, especially successful participants, learned to enter a meta-state that benefitted feedback and that showed to increase riPH later on.

One last note I would like to make with regards to this, is that while there is a hierarchical relationship between meta-states and states, there is yet another relationship between states to meta-states that is more rarely discussed. Meta-states are by far not independent from the states themselves even if it seems that they “control” them. If structural damage occurs to an actual brain state, this will have guaranteed impact in the meta-state organization due to a forced re-organization of connectivity (see: Caliendo et al., 2017; Grefkes & Fink, 2014). Furthermore, while studying meta-states might be of enormous significance for identification of hallucinations or other characteristics, certainly they are also bound to local dysfunctions known in certain diseases. A classical example would be in schizophrenia, where impairment at the level of the insula, likely drives also a change in meta-state^{††} related to the triple network model (Uddin, 2015).

5.2 Control of hallucinations

With Study II of this thesis, I made use of Study I findings, to establish a NF paradigm that would allow participants to achieve volitional control of the neural correlates of induced riPH. Three findings came out of this study. First, I identified the neural correlates of volitional control of the PH-network. Second, I showed that this training increased sensitivity to riPH post-training. Third, I demonstrated that for those who completed NF successfully and became sensitive to riPH there were lasting changes in brain dynamics.

A general advantage of NF studies, that also applies here, is the causality that can be established between the targeted neural marker and behavioral outcome. However, many times in hallucination research one confound remains that is that of targeting a trait versus state characteristic. In a first study of its kind, our participants achieved volitional control over a hallucination while it was being induced, effectively removing the confound of trait versus state. This has never been done in NF research of hallucinations, and is even very rarely done in NF research of behavioral traits (for an illustrative example in attention see: deBettencourt et al., 2015).

Another interesting consideration about my NF study was that participants learned to re-arrange the transition probability space to favor the PH-network, during the up-regulation condition of NF. They did re-arrangement, despite the NF feedback signal actually not containing direct information on transition probabilities. It could be argued that participants likely searched for different meta-states (i.e. transition configurations) which increased the received feedback. Now, for those that were successful, we don’t see a *permanence* of that meta-state post-training, but rather an after-effect that is translated into higher occurrences of the PH-network in the inducing condition. I would argue it indeed, would not have been plausible to have retained that meta-state, as it would have corresponded to a hallucination meta-state, and as I have mentioned before that is supposed to be a temporary configuration. Instead, we see an increase of the occurrences of the PH-network in the PH-inducing condition post-training. Given that Study I showed that it is not the occurrence but rather the

^{††} The reference does not explicitly mention meta-states. The reader should consider the concepts introduced in this subsection for the interpretation to meta-states

change in probabilities that is associated with riPH, I would further argue that what the protocol did was increase the sensitivity to riPH, by making it easier for participants to transition to that meta-state. In other words, the higher occurrence of the PH-network post-training most likely facilitates participants to transition to a meta-state that, in turn, makes them prone to PH-induction. Hence, the after effect can be seen as a lowering of the threshold to induce riPH, or a higher proneness to it.

5.3 Neuropsychiatry of Parkinson's Disease

The goal of study III was to provide an in-depth investigation on the neural correlates of hallucination progression in PD, and associate it with cognitive impairment, as well as our method to induce PH in patients with PD. A very important characteristic of this study was its unique stratification. To date, no previous work has investigated a “continuous” cohort of hallucination severity. Studies have compared patients with visual or minor hallucinations to those without hallucinations, but not across this complete progression. In doing so, we found two major findings, on two parts: behavior/cognitive and imaging.

Regarding behavior and cognitive findings, I found that sensitivity to riPH through our robotic procedure accompanied the increasing spectrum of severity. Patients with structured hallucinations are the most sensitive, followed by those with only minor hallucinations, and then by those without any hallucinations. This extended previous findings with this robotic system which had only been used in patients that had or not the specific minor hallucination (PH) the system is designed to induce (Bernasconi et al., 2021). A crucial aspect, however, is that sensitivity to riPH was modulated by an interaction between cognitive impairment and delay (a specific condition of the task) in a different manner across groups. In sum, sensitivity to riPH seems to increase for every delay in the same manner as cognitive impairment increase, for patients with structured hallucinations. However, for patients with no hallucinations, the sensitivity increases significantly more with cognitive impairment, the higher is the delay. The latter happens to the point that a cognitively preserved patient is not differently sensitive across delay but becomes distinctively sensitive if cognitively impaired. This could potentially indicate higher proneness to hallucinations as cognitive impairment increases, which has been suggested in the literature (Lenka et al., 2017), and could suggest that our robot might distinguish patients that are at the intersection of hallucination progression and cognitive decline.

The other major finding of this study, related to imaging, was the characterization of neural correlates of preserved cognition, lack of sensitivity to riPH, and no hallucinations in patients with PD. This was marked by less fluctuations of the visual network, observation of antagonistic activity between the DMN and both the visual network and an attentional network, as well as decreased antagonistic activity between the latter two. Again, these findings characterized preserved cognition and absence of neuropsychiatric symptoms. If we look at other studies with findings characterizing hallucination neural markers, such works actually found changes in the same networks reported here, however in the opposite direction, marking either structured hallucinations or cognitive impairment (e.g. hallucinations: Shine et al., 2015; cognitive impairment: Boord et al., 2017). Hence, I strongly believe that the uphold of these neural markers in PD are key to maintaining preserved function.

Jointly these two findings, show that it is possible to identify patients with no hallucinations and preserved cognition through the robot, and that these patients will have the neural characteristics

described above. At this point, considering the findings of my work, and those of others in the field, I would hypothesize the following: It is possible that the spectrum of sensitivity to riPH seen in our behavioral findings across patients with structured or no hallucinations and cognitively preserved or not, is in fact capturing the degeneration of these neural markers over the progression of the disease. It is likely that more data within the current study would already be enough to provide this evidence, but otherwise moving the experimental setup to the scanner could also answer this question.

5.4 Conclusion and outlook

The experimental work developed in this thesis was pioneering in the identification of temporal dynamics and control of hallucinations. I merged a state-of-the-art robotic system to induce a specific hallucination, and used novel methods to dissect the temporal structure of brain activity behind this hallucination. The pinnacle of this thesis was then using this knowledge to have participants manipulate the neural correlates of PH in real-time, which lead to an increase in sensitivity to riPH and lasting neural changes that matched the regulation mechanisms. I also identified neural correlates associated with preserved neuropsychiatric aspects of PD, and measured the various effects of our robotic task with a Bayesian meta-analysis.

I believe this thesis has already brought significant advances in the understanding of the underpinnings of hallucinations in both healthy individuals and PD. Study I and II provide sequential evidence of the importance of temporal dynamics, but in particular Study II provides significant causality between the PH-network activation and changes in transition probabilities and the sensitivity to PH-induction. They also surpass what others have done, as they are rid of the confounds of disease, and of the *constant trait* versus state dilemma presented previously. Then, Study II extended this, to volitional control of hallucinations. However, when considered with Study III, it brings potential clinical applications for antipsychotic therapies in PD, given that we have identified the correlates of preserved function as mentioned above, and shown these hallucinations to be controllable.

With this in mind, going forward there are two avenues of research that I believe should be explored. One which I will discuss first, comes from the results of Study I and some observations made in Study IV. It is a more conceptual avenue of research that pairs the analyses developed here, with experimental paradigms which would track measures of riPH during induction. The second one is a direct consequence of this thesis, in particular of my studies with neurofeedback, and in PD. That could be used for antipsychotic therapies in PD.

The first direction encompasses the study of meta-states (i.e. brain state interactions) in a paradigm capable of identifying induction of PH in short time intervals. This, would rely on pairing my findings and the dynamic connectivity methods used, with innovative experimental paradigms that could *implicitly* track riPH. One hypothesis would be to make use of numerosity estimations, as the original PH induction study (Blanke et al., 2014) saw participants overestimating the number of people with them in the experimental room with, while they were manipulating the robot in the condition PH is typically induced. There is work in the lab going in this experimental direction already (Albert et al., submitted). Translated to the MRI and paired it with the methods I am proposing, such work could reveal not just dynamics of brain states, but the underlying hierarchy of meta-states and states, and the relationship between the meta-states.

General Discussion

What I believe then would be the most consequential and promising avenue of research, would be to attempt to have PD patients reinstate the neural markers of preserved cognition and neuropsychiatric function through NF. This would be of considerable clinical significance as there is enough evidence showing the potential progress of hallucinations and their links to cognitive impairment (Lenka et al., 2017, 2019). Study II showed these neural correlates to be controllable, and hence reinforcing the neural markers identified in Study III could potentially delay the development of the disease. I want to stress out here, that this will not be a *cure*, as it would not stop the neurobiological progression of the disease. Nevertheless, there is enough evidence showing that personal coping strategies have an effect in reducing psychiatric progression (Diederich et al., 2003). Such an experiment would perhaps have to consider that PD patients might not be able to use the robot inside the MRI scanner, due to their motor comorbidities. Notwithstanding, NF could be done without fMRI – given that either way the goal would be to preserve brain function *not* associated with the induction/hallucination – and sessions with the robot could then be done outside the scanner, before and after training, to assess sensitivity to riPH. Crucially, if a strategy was found and confirmed to be retained in transfer sessions without NF, patients would be able to replicate that strategy at home, or potentially at the onset of hallucinations. The potential translational applications of this are therefore significant. For scientific research, a real-time fMRI NF experiment showing that successful up-regulation of neural markers related to preserved cognition and absent hallucinations leads to a delayed onset of hallucinations and cognitive decline, would provide causal evidence for the role of these networks in hallucinations in PD. However, for patients, it would be where most gain would be found. If such an experiment were successful, then patients could gain years of quality of life that would have otherwise been lost to psychiatric symptoms and cognitive decline.

As Anil Seth said, “reality, or at least our perception of it, is a controlled hallucination”. I would now end by adding that we can volitionally control the mechanisms leading up to these perceptions and by extension control our perception of reality

Herberto Dhanis

General Discussion

References

- Aarsland, D., Batzu, L., Halliday, G. M., Geurtsen, G. J., Ballard, C., Ray Chaudhuri, K., & Weintraub, D. (2021). Parkinson disease-associated cognitive impairment. *Nature Reviews Disease Primers*, 7(1), 47. <https://doi.org/10.1038/s41572-021-00280-3>
- Adams, R. A., Stephan, K. E., Brown, H. R., Frith, C. D., & Friston, K. J. (2013). The Computational Anatomy of Psychosis. *Frontiers in Psychiatry*, 4. <https://doi.org/10.3389/fpsy.2013.00047>
- Allefeld, C., Pütz, P., Kastner, K., & Wackermann, J. (2011). Flicker-light induced visual phenomena: Frequency dependence and specificity of whole percepts and percept features. *Consciousness and Cognition*, 20(4), 1344–1362. <https://doi.org/10.1016/j.concog.2010.10.026>
- Allen, L., O’Connell, A., & Kiermer, V. (2019). How can we ensure visibility and diversity in research contributions? How the Contributor Role Taxonomy (CRediT) is helping the shift from authorship to contributorship. *Learned Publishing*, 32(1), 71–74. <https://doi.org/10.1002/leap.1210>
- Allen, P., Larøi, F., McGuire, P. K., & Aleman, A. (2008). The hallucinating brain: A review of structural and functional neuroimaging studies of hallucinations. *Neuroscience & Biobehavioral Reviews*, 32(1), 175–191. <https://doi.org/10.1016/j.neubiorev.2007.07.012>
- Amano, K., Shibata, K., Kawato, M., Sasaki, Y., & Watanabe, T. (2016). Learning to Associate Orientation with Color in Early Visual Areas by Associative Decoded fMRI Neurofeedback. *Current Biology*, 26(14), 1861–1866. <https://doi.org/10.1016/j.cub.2016.05.014>
- Arns, M., Heinrich, H., & Strehl, U. (2014). Evaluation of neurofeedback in ADHD: The long and winding road. *Biological Psychology*, 95, 108–115. <https://doi.org/10.1016/j.biopsycho.2013.11.013>
- Arzy, S., Seeck, M., Ortigue, S., Spinelli, L., & Blanke, O. (2006). Induction of an illusory shadow person. *Nature*, 443(7109), 287–287. <https://doi.org/10.1038/443287a>
- Baggott, M. J., Siegrist, J. D., Galloway, G. P., Robertson, L. C., Coyle, J. R., & Mendelson, J. E. (2010). Investigating the Mechanisms of Hallucinogen-Induced Visions Using 3,4-Methylenedioxymphetamine (MDA): A Randomized Controlled Trial in Humans. *PLoS ONE*, 5(12), e14074. <https://doi.org/10.1371/journal.pone.0014074>
- Baker, J. T., Holmes, A. J., Masters, G. A., Yeo, B. T. T., Krienen, F., Buckner, R. L., & Öngür, D. (2014). Disruption of Cortical Association Networks in Schizophrenia and Psychotic Bipolar Disorder. *JAMA Psychiatry*, 71(2), 109. <https://doi.org/10.1001/jamapsychiatry.2013.3469>
- Bandettini, P. A., Wong, E. C., Hinks, R. S., Tikofsky, R. S., & Hyde, J. S. (1992). Time course EPI of human brain function during task activation. *Magnetic Resonance in Medicine*, 25(2), 390–397. <https://doi.org/10.1002/mrm.1910250220>
- Bauer, C. C. C., Okano, K., Ghosh, S. S., Lee, Y. J., Melero, H., Angeles, C. de los, Nestor, P. G., del Re, E. C., Northoff, G., Niznikiewicz, M. A., & Whitfield-Gabrieli, S. (2020). Real-time fMRI

References

- neurofeedback reduces auditory hallucinations and modulates resting state connectivity of involved brain regions: Part 2: Default mode network -preliminary evidence. *Psychiatry Research*, 284, 112770. <https://doi.org/10.1016/j.psychres.2020.112770>
- Beckmann, C. F., DeLuca, M., Devlin, J. T., & Smith, S. M. (2005). Investigations into resting-state connectivity using independent component analysis. *Philosophical Transactions of the Royal Society B: Biological Sciences*, 360(1457), 1001–1013. <https://doi.org/10.1098/rstb.2005.1634>
- Bejr-kasem, H., Pagonabarraga, J., Martínez-Horta, S., Sampedro, F., Marín-Lahoz, J., Horta-Barba, A., Aracil-Bolaños, I., Pérez-Pérez, J., Ángeles Botí, M., Campolongo, A., Izquierdo, C., Pascual-Sedano, B., Gómez-Ansón, B., & Kulisevsky, J. (2019). Disruption of the default mode network and its intrinsic functional connectivity underlies minor hallucinations in Parkinson's disease. *Movement Disorders*, 34(1), 78–86. <https://doi.org/10.1002/mds.27557>
- Bejr-kasem, H., Sampedro, F., Marín-Lahoz, J., Martínez-Horta, S., Pagonabarraga, J., & Kulisevsky, J. (2021). Minor hallucinations reflect early gray matter loss and predict subjective cognitive decline in Parkinson's disease. *European Journal of Neurology*, 28(2), 438–447. <https://doi.org/10.1111/ene.14576>
- Bernasconi, F., Blondiaux, E., Potheegadoo, J., Stripeikyte, G., Pagonabarraga, J., Bejr-Kasem, H., Bassolino, M., Akselrod, M., Martinez-Horta, S., Sampedro, F., Hara, M., Horvath, J., Franza, M., Konik, S., Bereau, M., Ghika, J.-A., Burkhard, P. R., Van De Ville, D., Faivre, N., ... Blanke, O. (2021). Robot-induced hallucinations in Parkinson's disease depend on altered sensorimotor processing in fronto-temporal network. *Science Translational Medicine*, 13(591), eabc8362. <https://doi.org/10.1126/scitranslmed.abc8362>
- Bernasconi, F., Blondiaux, E., Rognini, G., Dhanis, H., Jenni, L., Potheegadoo, J., Hara, M., & Blanke, O. (2022). Neuroscience robotics for controlled induction and real-time assessment of hallucinations. *Nature Protocols*. <https://doi.org/10.1038/s41596-022-00737-z>
- Betz, R. F., & Basset, D. S. (2017). Multi-scale brain networks. *NeuroImage*, 160, 73–83. <https://doi.org/10.1016/j.neuroimage.2016.11.006>
- Biswal, B., Zerrin Yetkin, F., Haughton, V. M., & Hyde, J. S. (1995). Functional connectivity in the motor cortex of resting human brain using echo-planar mri. *Magnetic Resonance in Medicine*, 34(4), 537–541. <https://doi.org/10.1002/mrm.1910340409>
- Blakemore, S.-J., Frith, C. D., & Wolpert, D. M. (1999). Spatio-Temporal Prediction Modulates the Perception of Self-Produced Stimuli. *Journal of Cognitive Neuroscience*, 11(5), 551–559. <https://doi.org/10.1162/089892999563607>
- Blakemore, S.-J., Wolpert, D. M., & Frith, C. D. (1998). Central cancellation of self-produced tickle sensation. *Nature Neuroscience*, 1(7), 635–640. <https://doi.org/10.1038/2870>
- Blakemore, S.-J., Wolpert, D. M., & Frith, C. D. (2002). Abnormalities in the awareness of action. *Trends in Cognitive Sciences*, 6(6), 237–242. [https://doi.org/10.1016/S1364-6613\(02\)01907-1](https://doi.org/10.1016/S1364-6613(02)01907-1)
- Blanke, O. (2012). Multisensory brain mechanisms of bodily self-consciousness. *Nature Reviews Neuroscience*, 13(8), 556–571. <https://doi.org/10.1038/nrn3292>

References

- Blanke, O., & Arzy, S. (2005). The Out-of-Body Experience: Disturbed Self-Processing at the Temporo-Parietal Junction. *The Neuroscientist*, 11(1), 16–24. <https://doi.org/10.1177/1073858404270885>
- Blanke, O., Landis, T., Spinelli, L., & Seeck, M. (2004). Out-of-body experience and autoscopia of neurological origin. *Brain*, 127(2), 243–258. <https://doi.org/10.1093/brain/awh040>
- Blanke, O., & Mohr, C. (2005). Out-of-body experience, heautoscopy, and autoscopic hallucination of neurological origin. *Brain Research Reviews*, 50(1), 184–199. <https://doi.org/10.1016/j.brainresrev.2005.05.008>
- Blanke, O., Ortigue, S., Coeytaux, A., Martory, M.-D., & Landis, T. (2003). Hearing of a Presence. *Neurocase*, 9(4), 329–339. <https://doi.org/10.1076/neur.9.4.329.15552>
- Blanke, O., Pozeg, P., Hara, M., Heydrich, L., Serino, A., Yamamoto, A., Higuchi, T., Salomon, R., Seeck, M., Landis, T., Arzy, S., Herbelin, B., Bleuler, H., & Rognini, G. (2014). Neurological and Robot-Controlled Induction of an Apparition. *Current Biology*, 24(22), 2681–2686. <https://doi.org/10.1016/j.cub.2014.09.049>
- Blondiaux, E. (2020). *Hallucination Engineering: From Neuroscience Robotics to Hallucinations in Health and Disease*. École Polytechnique Fédérale de Lausanne.
- Blondiaux, E., Heydrich, L., & Blanke, O. (2021). Common and distinct brain networks of autoscopic phenomena. *NeuroImage: Clinical*, 30, 102612. <https://doi.org/10.1016/j.nicl.2021.102612>
- Bluhm, R. L., Miller, J., Lanius, R. A., Osuch, E. A., Boksman, K., Neufeld, R., Theberge, J., Schaefer, B., & Williamson, P. (2007). Spontaneous Low-Frequency Fluctuations in the BOLD Signal in Schizophrenic Patients: Anomalies in the Default Network. *Schizophrenia Bulletin*, 33(4), 1004–1012. <https://doi.org/10.1093/schbul/sbm052>
- Boas, D. A., Dale, A. M., & Franceschini, M. A. (2004). Diffuse optical imaging of brain activation: Approaches to optimizing image sensitivity, resolution, and accuracy. *NeuroImage*, 23, S275–S288. <https://doi.org/10.1016/j.neuroimage.2004.07.011>
- Boes, A. D., Prasad, S., Liu, H., Liu, Q., Pascual-Leone, A., Caviness, V. S., & Fox, M. D. (2015). Network localization of neurological symptoms from focal brain lesions. *Brain*, 138(10), 3061–3075. <https://doi.org/10.1093/brain/awv228>
- Boord, P., Madhyastha, T. M., Askren, M. K., & Grabowski, T. J. (2017). Executive attention networks show altered relationship with default mode network in PD. *NeuroImage: Clinical*, 13, 1–8. <https://doi.org/10.1016/j.nicl.2016.11.004>
- Botvinick, M., & Cohen, J. (1998). Rubber hands ‘feel’ touch that eyes see. *Nature*, 391(6669), 756–756. <https://doi.org/10.1038/35784>
- Brugger, P., Regard, M., & Landis, T. (1997). Illusory Reduplication of One’s Own Body: Phenomenology and Classification of Autoscopic Phenomena. *Cognitive Neuropsychiatry*, 2(1), 19–38. <https://doi.org/10.1080/135468097396397>
- Bubic, A., von Cramon, D. Y., & Schubotz, R. (2010). Prediction, cognition and the brain. *Frontiers in Human Neuroscience*. <https://doi.org/10.3389/fnhum.2010.00025>

References

- Buckner, R. L., Andrews-Hanna, J. R., & Schacter, D. L. (2008). *The Brain's Default Network: Anatomy, Function, and Relevance to Disease*. *Annals of the New York Academy of Sciences*, 1124(1), 1–38. <https://doi.org/10.1196/annals.1440.011>
- Cachia, A., Cury, C., Brunelin, J., Plaze, M., Delmaire, C., Oppenheim, C., Medjkane, F., Thomas, P., & Jardri, R. (2020). Deviations in early hippocampus development contribute to visual hallucinations in schizophrenia. *Translational Psychiatry*, 10(1), 102. <https://doi.org/10.1038/s41398-020-0779-9>
- Caliandro, P., Vecchio, F., Miraglia, F., Reale, G., Della Marca, G., La Torre, G., Lacidogna, G., Iacovelli, C., Padua, L., Bramanti, P., & Rossini, P. M. (2017). Small-World Characteristics of Cortical Connectivity Changes in Acute Stroke. *Neurorehabilitation and Neural Repair*, 31(1), 81–94. <https://doi.org/10.1177/1545968316662525>
- Camchong, J., MacDonald, A. W., Bell, C., Mueller, B. A., & Lim, K. O. (2011). Altered Functional and Anatomical Connectivity in Schizophrenia. *Schizophrenia Bulletin*, 37(3), 640–650. <https://doi.org/10.1093/schbul/sbp131>
- Caria, A., Sitaram, R., & Birbaumer, N. (2012). Real-Time fMRI: A Tool for Local Brain Regulation. *The Neuroscientist*, 18(5), 487–501. <https://doi.org/10.1177/1073858411407205>
- Caria, A., Veit, R., Sitaram, R., Lotze, M., Weiskopf, N., Grodd, W., & Birbaumer, N. (2007). Regulation of anterior insular cortex activity using real-time fMRI. *NeuroImage*, 35(3), 1238–1246. <https://doi.org/10.1016/j.neuroimage.2007.01.018>
- Case, L. K., Solcà, M., Blanke, O., & Faivre, N. (2020). Disorders of body representation. In *Multisensory Perception* (pp. 401–422). Elsevier. <https://doi.org/10.1016/B978-0-12-812492-5.00018-8>
- Cassidy, C. M., Balsam, P. D., Weinstein, J. J., Rosengard, R. J., Slifstein, M., Daw, N. D., Abi-Dargham, A., & Horga, G. (2018). A Perceptual Inference Mechanism for Hallucinations Linked to Striatal Dopamine. *Current Biology*, 28(4), 503–514.e4. <https://doi.org/10.1016/j.cub.2017.12.059>
- Chambon, V., Wenke, D., Fleming, S. M., Prinz, W., & Haggard, P. (2013). An Online Neural Substrate for Sense of Agency. *Cerebral Cortex*, 23(5), 1031–1037. <https://doi.org/10.1093/cercor/bhs059>
- Chan, D., & Rossor, M. N. (2002). “—But who is that on the other side of you?” Extracampine hallucinations revisited. *The Lancet*, 360(9350), 2064–2066. [https://doi.org/10.1016/S0140-6736\(02\)11998-2](https://doi.org/10.1016/S0140-6736(02)11998-2)
- Chang, C., & Glover, G. H. (2010). Time–frequency dynamics of resting-state brain connectivity measured with fMRI. *NeuroImage*, 50(1), 81–98. <https://doi.org/10.1016/j.neuroimage.2009.12.011>
- Chang, X., Collin, G., Xi, Y., Cui, L., Scholtens, L. H., Sommer, I. E., Wang, H., Yin, H., Kahn, R. S., & van den Heuvel, M. P. (2017). Resting-state functional connectivity in medication-naïve schizophrenia patients with and without auditory verbal hallucinations: A preliminary report. *Schizophrenia Research*, 188, 75–81. <https://doi.org/10.1016/j.schres.2017.01.024>

References

- Chen, J. E., Chang, C., Greicius, M. D., & Glover, G. H. (2015). Introducing co-activation pattern metrics to quantify spontaneous brain network dynamics. *NeuroImage*, 111, 476–488. <https://doi.org/10.1016/j.neuroimage.2015.01.057>
- Chen, T., Cai, W., Ryali, S., Supekar, K., & Menon, V. (2016). Distinct Global Brain Dynamics and Spatiotemporal Organization of the Salience Network. *PLOS Biology*, 14(6), e1002469. <https://doi.org/10.1371/journal.pbio.1002469>
- Cole, M. W., Anticevic, A., Repovs, G., & Barch, D. (2011). Variable Global Dysconnectivity and Individual Differences in Schizophrenia. *Biological Psychiatry*, 70(1), 43–50. <https://doi.org/10.1016/j.biopsych.2011.02.010>
- Colonius, H., & Diederich, A. (2004). Multisensory Interaction in Saccadic Reaction Time: A Time-Window-of-Integration Model. *Journal of Cognitive Neuroscience*, 16(6), 1000–1009. <https://doi.org/10.1162/0898929041502733>
- Corlett, P. R., Horga, G., Fletcher, P. C., Alderson-Day, B., Schmack, K., & Powers, A. R. (2019). Hallucinations and Strong Priors. *Trends in Cognitive Sciences*, 23(2), 114–127. <https://doi.org/10.1016/j.tics.2018.12.001>
- Damasio, H., Grabowski, T., Frank, R., Galaburda, A. M., & Damasio, A. R. (1994). The Return of Phineas Gage: Clues About the Brain from the Skull of a Famous Patient. *Science*, 264(5162), 1102–1105. <https://doi.org/10.1126/science.8178168>
- Damoiseaux, J. S., Rombouts, S. A. R. B., Barkhof, F., Scheltens, P., Stam, C. J., Smith, S. M., & Beckmann, C. F. (2006). Consistent resting-state networks across healthy subjects. *Proceedings of the National Academy of Sciences*, 103(37), 13848–13853. <https://doi.org/10.1073/pnas.0601417103>
- De Haan, E. H., Nys, G. M., van Zandvoort, M. J., & Ramsey, N. F. (2007). The physiological basis of visual hallucinations after damage to the primary visual cortex. *NeuroReport*, 18(11), 1177–1180. <https://doi.org/10.1097/WNR.0b013e32820049d3>
- de Lange, F. P., Heilbron, M., & Kok, P. (2018). How Do Expectations Shape Perception? *Trends in Cognitive Sciences*, 22(9), 764–779. <https://doi.org/10.1016/j.tics.2018.06.002>
- deBettencourt, M. T., Cohen, J. D., Lee, R. F., Norman, K. A., & Turk-Browne, N. B. (2015). Closed-loop training of attention with real-time brain imaging. *Nature Neuroscience*, 18(3), 470–475. <https://doi.org/10.1038/nn.3940>
- deCharms, R. C., Christoff, K., Glover, G. H., Pauly, J. M., Whitfield, S., & Gabrieli, J. D. E. (2004). Learned regulation of spatially localized brain activation using real-time fMRI. *NeuroImage*, 21(1), 436–443. <https://doi.org/10.1016/j.neuroimage.2003.08.041>
- deCharms, R. C., Maeda, F., Glover, G. H., Ludlow, D., Pauly, J. M., Soneji, D., Gabrieli, J. D. E., & Mackey, S. C. (2005). Control over brain activation and pain learned by using real-time functional MRI. *Proceedings of the National Academy of Sciences*, 102(51), 18626–18631. <https://doi.org/10.1073/pnas.0505210102>
- Diederich, N. J., Fénelon, G., Stebbins, G., & Goetz, C. G. (2009). Hallucinations in Parkinson disease. *Nature Reviews Neurology*, 5(6), 331–342. <https://doi.org/10.1038/nrneuro.2009.62>

References

- Diederich, N. J., Pieri, V., & Goetz, C. G. (2003). Coping strategies for visual hallucinations in Parkinson's disease. *Movement Disorders*, 18(7), 831–832. <https://doi.org/10.1002/mds.10450>
- Dronkers, N. F., Plaisant, O., Iba-Zizen, M. T., & Cabanis, E. A. (2007). Paul Broca's historic cases: High resolution MR imaging of the brains of Leborgne and Lelong. *Brain*, 130(5), 1432–1441. <https://doi.org/10.1093/brain/awm042>
- Eagleman, D. M. (2001). Visual illusions and neurobiology. *Nature Reviews Neuroscience*, 2(12), 920–926. <https://doi.org/10.1038/35104092>
- Ehrsson, H. H., Spence, C., & Passingham, R. E. (2004). That's My Hand! Activity in Premotor Cortex Reflects Feeling of Ownership of a Limb. *Science*, 305(5685), 875–877. <https://doi.org/10.1126/science.1097011>
- Ehrsson, H. H., Wiech, K., Weiskopf, N., Dolan, R. J., & Passingham, R. E. (2007). Threatening a rubber hand that you feel is yours elicits a cortical anxiety response. *Proceedings of the National Academy of Sciences*, 104(23), 9828–9833. <https://doi.org/10.1073/pnas.0610011104>
- Factor, S. A., Feustel, P. J., Friedman, J. H., Comella, C. L., Goetz, C. G., Kurlan, R., Parsa, M., Pfeiffer, R., & the Parkinson Study Group. (2003). Longitudinal outcome of Parkinson's disease patients with psychosis. *Neurology*, 60(11), 1756–1761. <https://doi.org/10.1212/01.WNL.0000068010.82167.CF>
- Farrer, C., & Franck, N. (2007). Self-Monitoring in Schizophrenia. *Current Psychiatry Reviews*, 3(4), 243–251. <https://doi.org/10.2174/157340007782408897>
- Farrer, C., Franck, N., Paillard, J., & Jeannerod, M. (2003). The role of proprioception in action recognition. *Consciousness and Cognition*, 12(4), 609–619. [https://doi.org/10.1016/S1053-8100\(03\)00047-3](https://doi.org/10.1016/S1053-8100(03)00047-3)
- Farrer, C., & Frith, C. D. (2002). Experiencing Oneself vs Another Person as Being the Cause of an Action: The Neural Correlates of the Experience of Agency. *NeuroImage*, 15(3), 596–603. <https://doi.org/10.1006/nimg.2001.1009>
- Fenelon, G. (2000). Hallucinations in Parkinson's disease: Prevalence, phenomenology and risk factors. *Brain*, 123(4), 733–745. <https://doi.org/10.1093/brain/123.4.733>
- Fenelon, G., Goetz, C. G., & Karenberg, A. (2006). Hallucinations in Parkinson disease in the prelevodopa era. *Neurology*, 66(1), 93–98. <https://doi.org/10.1212/01.wnl.0000191325.31068.c4>
- Fenelon, G., Soulas, T., de Langavant, L. C., Trinkler, I., & Bachoud-Levi, A.-C. (2011). Feeling of presence in Parkinson's disease. *Journal of Neurology, Neurosurgery & Psychiatry*, 82(11), 1219–1224. <https://doi.org/10.1136/jnnp.2010.234799>
- ffytche, D. H., Creese, B., Politis, M., Chaudhuri, K. R., Weintraub, D., Ballard, C., & Aarsland, D. (2017). The psychosis spectrum in Parkinson disease. *Nature Reviews Neurology*, 13(2), 81–95. <https://doi.org/10.1038/nrneurol.2016.200>
- Finn, E. S., Shen, X., Scheinost, D., Rosenberg, M. D., Huang, J., Chun, M. M., Papademetris, X., & Constable, R. T. (2015). Functional connectome fingerprinting: Identifying individuals using

References

- patterns of brain connectivity. *Nature Neuroscience*, 18(11), 1664–1671. <https://doi.org/10.1038/nn.4135>
- Flint, A. C., Loh, J. P., & Brust, J. C. M. (2005). Vivid visual hallucinations from occipital lobe infarction. *Neurology*, 65(5), 756–756. <https://doi.org/10.1212/01.wnl.0000180350.58985.5f>
- Ford, J. M., Mathalon, D. H., Heinks, T., Kalba, S., Faustman, W. O., & Roth, W. T. (2001). Neurophysiological Evidence of Corollary Discharge Dysfunction in Schizophrenia. *American Journal of Psychiatry*, 158(12), 2069–2071. <https://doi.org/10.1176/appi.ajp.158.12.2069>
- Ford, J. M., Mathalon, D. H., Whitfield, S., Faustman, W. O., & Roth, W. T. (2002). Reduced communication between frontal and temporal lobes during talking in schizophrenia. *Biological Psychiatry*, 51(6), 485–492. [https://doi.org/10.1016/S0006-3223\(01\)01335-X](https://doi.org/10.1016/S0006-3223(01)01335-X)
- Ford, J. M., Roach, B. J., Faustman, W. O., & Mathalon, D. H. (2007). Synch Before You Speak: Auditory Hallucinations in Schizophrenia. *Am J Psychiatry*, 9.
- Fornito, A., Yoon, J., Zalesky, A., Bullmore, E. T., & Carter, C. S. (2011). General and Specific Functional Connectivity Disturbances in First-Episode Schizophrenia During Cognitive Control Performance. *Biological Psychiatry*, 70(1), 64–72. <https://doi.org/10.1016/j.biopsych.2011.02.019>
- Fox, M. D., Corbetta, M., Snyder, A. Z., Vincent, J. L., & Raichle, M. E. (2006). Spontaneous neuronal activity distinguishes human dorsal and ventral attention systems. *Proceedings of the National Academy of Sciences*, 103(26), 10046–10051. <https://doi.org/10.1073/pnas.0604187103>
- Fox, M. D., & Raichle, M. E. (2007). Spontaneous fluctuations in brain activity observed with functional magnetic resonance imaging. *Nature Reviews Neuroscience*, 8(9), 700–711. <https://doi.org/10.1038/nrn2201>
- Fox, M. D., Snyder, A. Z., Vincent, J. L., Corbetta, M., Van Essen, D. C., & Raichle, M. E. (2005). The human brain is intrinsically organized into dynamic, anticorrelated functional networks. *Proceedings of the National Academy of Sciences*, 102(27), 9673–9678. <https://doi.org/10.1073/pnas.0504136102>
- Friston, K., Brown, H. R., Siemerkus, J., & Stephan, K. E. (2016). The dysconnection hypothesis (2016). *Schizophrenia Research*, 176(2–3), 83–94. <https://doi.org/10.1016/j.schres.2016.07.014>
- Friston, K. J. (2011). Functional and Effective Connectivity: A Review. *Brain Connectivity*, 1(1), 13–36. <https://doi.org/10.1089/brain.2011.0008>
- Friston, K. J., Holmes, A. P., Worsley, K. J., Poline, J.-P., Frith, C. D., & Frackowiak, R. S. J. (1994). Statistical parametric maps in functional imaging: A general linear approach. *Human Brain Mapping*, 2(4), 189–210. <https://doi.org/10.1002/hbm.460020402>
- Friston, K. J., & Stephan, K. E. (2007). Free-energy and the brain. *Synthese*, 159(3), 417–458. <https://doi.org/10.1007/s11229-007-9237-y>
- Frith, C. D. (1987). The positive and negative symptoms of schizophrenia reflect impairments in the perception and initiation of action. *Psychological Medicine*, 17(3), 631–648. <https://doi.org/10.1017/S0033291700025873>

References

- Geiger, J. (2009). *The third man factor*. Hachette Books.
- Ghahramani, Z., Wolpert, D. M., & Jordan, M. I. (1997). Computational models of sensorimotor integration. In *Advances in Psychology* (Vol. 119, pp. 117–147). Elsevier. [https://doi.org/10.1016/S0166-4115\(97\)80006-4](https://doi.org/10.1016/S0166-4115(97)80006-4)
- Goetz, C. G., Fan, W., Leurgans, S., Bernard, B., & Stebbins, G. T. (2006). The Malignant Course of “Benign Hallucinations” in Parkinson Disease. *Archives of Neurology*, 63(5), 713. <https://doi.org/10.1001/archneur.63.5.713>
- Gore, J. C. (2003). Principles and practice of functional MRI of the human brain. *Journal of Clinical Investigation*, 112(1), 4–9. <https://doi.org/10.1172/JCI200319010>
- Goulden, N., Khusnulina, A., Davis, N. J., Bracewell, R. M., Bokde, A. L., McNulty, J. P., & Mullins, P. G. (2014). The salience network is responsible for switching between the default mode network and the central executive network: Replication from DCM. *NeuroImage*, 99, 180–190. <https://doi.org/10.1016/j.neuroimage.2014.05.052>
- Grefkes, C., & Fink, G. R. (2014). Connectivity-based approaches in stroke and recovery of function. *The Lancet Neurology*, 13(2), 206–216. [https://doi.org/10.1016/S1474-4422\(13\)70264-3](https://doi.org/10.1016/S1474-4422(13)70264-3)
- Haggard, P., & Chambon, V. (2012). Sense of agency. *Current Biology*, 22(10), R390–R392. <https://doi.org/10.1016/j.cub.2012.02.040>
- Ham, T., Leff, A., de Boissezon, X., Joffe, A., & Sharp, D. J. (2013). Cognitive Control and the Salience Network: An Investigation of Error Processing and Effective Connectivity. *Journal of Neuroscience*, 33(16), 7091–7098. <https://doi.org/10.1523/JNEUROSCI.4692-12.2013>
- Hara, M., Salomon, R., van der Zwaag, W., Kober, T., Rognini, G., Nabae, H., Yamamoto, A., Blanke, O., & Higuchi, T. (2014). A novel manipulation method of human body ownership using an fMRI-compatible master–slave system. *Journal of Neuroscience Methods*, 235, 25–34. <https://doi.org/10.1016/j.jneumeth.2014.05.038>
- Haxby, J. V., Gobbini, M. I., Furey, M. L., Ishai, A., Schouten, J. L., & Pietrini, P. (2001). Distributed and Overlapping Representations of Faces and Objects in Ventral Temporal Cortex. *Science*, 293(5539), 2425–2430. <https://doi.org/10.1126/science.1063736>
- Haxby, J. V., Guntupalli, J. S., Connolly, A. C., Halchenko, Y. O., Conroy, B. R., Gobbini, M. I., Hanke, M., & Ramadge, P. J. (2011). A Common, High-Dimensional Model of the Representational Space in Human Ventral Temporal Cortex. *Neuron*, 72(2), 404–416. <https://doi.org/10.1016/j.neuron.2011.08.026>
- Hayes, J., & Leudar, I. (2016). Experiences of continued presence: On the practical consequences of ‘hallucinations’ in bereavement. *Psychology and Psychotherapy: Theory, Research and Practice*, 89(2), 194–210. <https://doi.org/10.1111/papt.12067>
- Heeger, D. J., & Ress, D. (2002). What does fMRI tell us about neuronal activity? *Nature Reviews Neuroscience*, 3(2), 142–151. <https://doi.org/10.1038/nrn730>
- Hely, M. A., Reid, W. G. J., Adena, M. A., Halliday, G. M., & Morris, J. G. L. (2008). The Sydney multicenter study of Parkinson’s disease: The inevitability of dementia at 20 years: Twenty

References

- Year Sydney Parkinson's Study. *Movement Disorders*, 23(6), 837–844. <https://doi.org/10.1002/mds.21956>
- Hutchison, R. M., Womelsdorf, T., Gati, J. S., Everling, S., & Menon, R. S. (2013). Resting-state networks show dynamic functional connectivity in awake humans and anesthetized macaques: Dynamic Functional Connectivity. *Human Brain Mapping*, 34(9), 2154–2177. <https://doi.org/10.1002/hbm.22058>
- Ibarretxe-Bilbao, N., Ramirez-Ruiz, B., Junque, C., Marti, M. J., Valldeoriola, F., Bargallo, N., Juanes, S., & Tolosa, E. (2010). Differential progression of brain atrophy in Parkinson's disease with and without visual hallucinations. *Journal of Neurology, Neurosurgery & Psychiatry*, 81(6), 650–657. <https://doi.org/10.1136/jnnp.2009.179655>
- Jardri, R., Pouchet, A., Pins, D., & Thomas, P. (2011). Cortical Activations During Auditory Verbal Hallucinations in Schizophrenia: A Coordinate-Based Meta-Analysis. *American Journal of Psychiatry*, 168(1), 73–81. <https://doi.org/10.1176/appi.ajp.2010.09101522>
- Jardri, R., Thomas, P., Delmaire, C., Delion, P., & Pins, D. (2013). The Neurodynamic Organization of Modality-Dependent Hallucinations. *Cerebral Cortex*, 23(5), 1108–1117. <https://doi.org/10.1093/cercor/bhs082>
- Kaiser, D., Quek, G. L., Cichy, R. M., & Peelen, M. V. (2019). Object Vision in a Structured World. *Trends in Cognitive Sciences*, 23(8), 672–685. <https://doi.org/10.1016/j.tics.2019.04.013>
- Kalia, L. V., & Lang, A. E. (2015). Parkinson's disease. *The Lancet*, 386(9996), 896–912. [https://doi.org/10.1016/S0140-6736\(14\)61393-3](https://doi.org/10.1016/S0140-6736(14)61393-3)
- Karbasforoushan, H., & Woodward, N. D. (2012). Resting-State Networks in Schizophrenia. *Current Topics in Medical Chemistry*, 12, 2404–2414.
- Koizumi, A., Amano, K., Cortese, A., Shibata, K., Yoshida, W., Seymour, B., Kawato, M., & Lau, H. (2017). Fear reduction without fear through reinforcement of neural activity that bypasses conscious exposure. *Nature Human Behaviour*, 1(1), 0006. <https://doi.org/10.1038/s41562-016-0006>
- Kok, P., Brouwer, G. J., van Gerven, M. A. J., & de Lange, F. P. (2013). Prior Expectations Bias Sensory Representations in Visual Cortex. *Journal of Neuroscience*, 33(41), 16275–16284. <https://doi.org/10.1523/JNEUROSCI.0742-13.2013>
- Kok, P., Jehee, J. F. M., & de Lange, F. P. (2012). Less Is More: Expectation Sharpens Representations in the Primary Visual Cortex. *Neuron*, 75(2), 265–270. <https://doi.org/10.1016/j.neuron.2012.04.034>
- Koush, Y., Rosa, M. J., Robineau, F., Heinen, K., W. Rieger, S., Weiskopf, N., Vuilleumier, P., Van De Ville, D., & Scharnowski, F. (2013). Connectivity-based neurofeedback: Dynamic causal modeling for real-time fMRI. *NeuroImage*, 81, 422–430. <https://doi.org/10.1016/j.neuroimage.2013.05.010>
- Krestel, H. (2013). Language and brain: Historical introduction to models of language and aphasia. *Schweizer Archiv Für Neurologie Und Psychiatrie*, 164(08), 262–266. <https://doi.org/10.4414/sanp.2013.00211>

References

- Kucyi, A., & Davis, K. D. (2014). Dynamic functional connectivity of the default mode network tracks daydreaming. *NeuroImage*, 100, 471–480. <https://doi.org/10.1016/j.neuroimage.2014.06.044>
- Kwong, K. K., Belliveau, J. W., Chesler, D. A., Goldberg, I. E., Weisskoff, R. M., Poncelet, B. P., Kennedy, D. N., Hoppel, B. E., Cohen, M. S., & Turner, R. (1992). Dynamic magnetic resonance imaging of human brain activity during primary sensory stimulation. *Proceedings of the National Academy of Sciences*, 89(12), 5675–5679. <https://doi.org/10.1073/pnas.89.12.5675>
- Lameka, K., Farwell, M. D., & Ichise, M. (2016). Positron Emission Tomography. In *Handbook of Clinical Neurology* (Vol. 135, pp. 209–227). Elsevier. <https://doi.org/10.1016/B978-0-444-53485-9.00011-8>
- Larøi, F., Bless, J. J., Laloyaux, J., Kråkvik, B., Vedul-Kjelsås, E., Kalhovde, A. M., Hirnstein, M., & Hugdahl, K. (2019). An epidemiological study on the prevalence of hallucinations in a general-population sample: Effects of age and sensory modality. *Psychiatry Research*, 272, 707–714. <https://doi.org/10.1016/j.psychres.2019.01.003>
- Lefebvre, S., Demeulemeester, M., Leroy, A., Delmaire, C., Lopes, R., Pins, D., Thomas, P., & Jardri, R. (2016). Network dynamics during the different stages of hallucinations in schizophrenia: Network Dynamics During Hallucinations. *Human Brain Mapping*, 37(7), 2571–2586. <https://doi.org/10.1002/hbm.23197>
- Lenggenhager, B., Tadi, T., Metzinger, T., & Blanke, O. (2007). Video Ergo Sum: Manipulating Bodily Self-Consciousness. *Science*, 317(5841), 1096–1099. <https://doi.org/10.1126/science.1143439>
- Lenka, A., George, L., Arumugham, S. S., Hegde, S., Reddy, V., Kamble, N., Yadav, R., & Pal, P. K. (2017). Predictors of onset of psychosis in patients with Parkinson’s disease: Who gets it early? *Parkinsonism & Related Disorders*, 44, 91–94. <https://doi.org/10.1016/j.parkreldis.2017.09.015>
- Lenka, A., Pagonabarraga, J., Pal, P. K., Bejr-Kasem, H., & Kulisevsky, J. (2019). Minor hallucinations in Parkinson disease: A subtle symptom with major clinical implications. *Neurology*, 93(6), 259–266. <https://doi.org/10.1212/WNL.00000000000007913>
- Linscott, R. J., & van Os, J. (2010). Systematic Reviews of Categorical Versus Continuum Models in Psychosis: Evidence for Discontinuous Subpopulations Underlying a Psychometric Continuum. Implications for DSM-V, DSM-VI, and DSM-VII. *Annual Review of Clinical Psychology*, 6(1), 391–419. <https://doi.org/10.1146/annurev.clinpsy.032408.153506>
- Liu, X., & Duyn, J. H. (2013). Time-varying functional network information extracted from brief instances of spontaneous brain activity. *Proceedings of the National Academy of Sciences*, 110(11), 4392–4397. <https://doi.org/10.1073/pnas.1216856110>
- Liu, X., Zhang, N., Chang, C., & Duyn, J. H. (2018). Co-activation patterns in resting-state fMRI signals. *NeuroImage*, 180, 485–494. <https://doi.org/10.1016/j.neuroimage.2018.01.041>
- Llebaria, G., Pagonabarraga, J., Martínez-Corral, M., García-Sánchez, C., Pascual-Sedano, B., Gironell, A., & Kulisevsky, J. (2010). Neuropsychological correlates of mild to severe hallucinations in

References

- Parkinson's disease: Neuropsychology of VH in Parkinson's Disease. *Movement Disorders*, 25(16), 2785–2791. <https://doi.org/10.1002/mds.23411>
- Llorca, P. M., Pereira, B., Jardri, R., Chereau-Boudet, I., Brousse, G., Misdrahi, D., Fénelon, G., Tronche, A.-M., Schwan, R., Lançon, C., Marques, A., Ulla, M., Derost, P., Debilly, B., Durif, F., & de Chazeron, I. (2016). Hallucinations in schizophrenia and Parkinson's disease: An analysis of sensory modalities involved and the repercussion on patients. *Scientific Reports*, 6(1), 38152. <https://doi.org/10.1038/srep38152>
- Logothetis, N. K., Pauls, J., Augath, M., Trinath, T., & Oeltermann, A. (2001). *Neurophysiological investigation of the basis of the fMRI signal*. 412, 8.
- Lopes da Silva, F. (2013). EEG and MEG: Relevance to Neuroscience. *Neuron*, 80(5), 1112–1128. <https://doi.org/10.1016/j.neuron.2013.10.017>
- Maijer, K., Begemann, M. J. H., Palmen, S. J. M. C., Leucht, S., & Sommer, I. E. C. (2018). Auditory hallucinations across the lifespan: A systematic review and meta-analysis. *Psychological Medicine*, 48(6), 879–888. <https://doi.org/10.1017/S0033291717002367>
- Manford, M., & Andermann, F. (1998). Complex visual hallucinations. Clinical and neurobiological insights. *Brain*, 121(10), 1819–1840. <https://doi.org/10.1093/brain/121.10.1819>
- Manoliu, A., Riedl, V., Zherdin, A., Mühlau, M., Schwerthöffer, D., Scherr, M., Peters, H., Zimmer, C., Förstl, H., Bäuml, J., Wohlschläger, A. M., & Sorg, C. (2014). Aberrant Dependence of Default Mode/Central Executive Network Interactions on Anterior Insular Salience Network Activity in Schizophrenia. *Schizophrenia Bulletin*, 40(2), 428–437. <https://doi.org/10.1093/schbul/sbt037>
- McKeith, I. G., Boeve, B. F., Dickson, D. W., Halliday, G., Taylor, J.-P., Weintraub, D., Aarsland, D., Galvin, J., Attems, J., Ballard, C. G., Bayston, A., Beach, T. G., Blanc, F., Bohnen, N., Bonanni, L., Bras, J., Brundin, P., Burn, D., Chen-Plotkin, A., ... Kosaka, K. (2017). Diagnosis and management of dementia with Lewy bodies: Fourth consensus report of the DLB Consortium. *Neurology*, 89(1), 88–100. <https://doi.org/10.1212/WNL.0000000000004058>
- Megumi, F., Yamashita, A., Kawato, M., & Imamizu, H. (2015). Functional MRI neurofeedback training on connectivity between two regions induces long-lasting changes in intrinsic functional network. *Frontiers in Human Neuroscience*, 9. <https://doi.org/10.3389/fnhum.2015.00160>
- Menezes, N. M., Arenovich, T., & Zipursky, R. B. (2006). A systematic review of longitudinal outcome studies of first-episode psychosis. *Psychological Medicine*, 36(10), 1349–1362. <https://doi.org/10.1017/S0033291706007951>
- Menon, V. (2011). Large-scale brain networks and psychopathology: A unifying triple network model. *Trends in Cognitive Sciences*, 15(10), 483–506. <https://doi.org/10.1016/j.tics.2011.08.003>
- Menon, V., & Uddin, L. Q. (2010). Saliency, switching, attention and control: A network model of insula function. *Brain Structure and Function*, 214(5–6), 655–667. <https://doi.org/10.1007/s00429-010-0262-0>
- Messner, R. (2016). *Naked Mountain: Nanga Parbat, Brother, Death, Solitude*. Crowood.

References

- Miall, R. C., & Wolpert, D. M. (1996). Forward Models for Physiological Motor Control. *Neural Networks*, 9(8), 1265–1279. [https://doi.org/10.1016/S0893-6080\(96\)00035-4](https://doi.org/10.1016/S0893-6080(96)00035-4)
- Millan, M. J., Andrieux, A., Bartzokis, G., Cadenhead, K., Dazzan, P., Fusar-Poli, P., Gallinat, J., Giedd, J., Grayson, D. R., Heinrichs, M., Kahn, R., Krebs, M.-O., Leboyer, M., Lewis, D., Marin, O., Marin, P., Meyer-Lindenberg, A., McGorry, P., McGuire, P., ... Weinberger, D. (2016). Altering the course of schizophrenia: Progress and perspectives. *Nature Reviews Drug Discovery*, 15(7), 485–515. <https://doi.org/10.1038/nrd.2016.28>
- Mlakar, J., Jensterle, J., & Frith, C. D. (1994). Central monitoring deficiency and schizophrenic symptoms. *Psychological Medicine*, 24(3), 557–564. <https://doi.org/10.1017/S0033291700027719>
- Moran, L. V., Tagamets, M. A., Sampath, H., O'Donnell, A., Stein, E. A., Kochunov, P., & Hong, L. E. (2013). Disruption of Anterior Insula Modulation of Large-Scale Brain Networks in Schizophrenia. *Biological Psychiatry*, 74(6), 467–474. <https://doi.org/10.1016/j.biopsych.2013.02.029>
- Morgante, L., Colosimo, C., Antonini, A., Marconi, R., Meco, G., Pederzoli, M., Pontieri, F. E., Cicarelli, G., Abbruzzese, G., Zappulla, S., Ramat, S., Manfredi, M., Bottacchi, E., Abrignani, M., Berardelli, A., Cozzolino, A., Paradiso, C., De Gaspari, D., Morgante, F., ... on behalf of the PRIAMO Study Group. (2012). Psychosis associated to Parkinson's disease in the early stages: Relevance of cognitive decline and depression. *Journal of Neurology, Neurosurgery & Psychiatry*, 83(1), 76–82. <https://doi.org/10.1136/jnnp-2011-300043>
- Mucha, P. J., Richardson, T., Macon, K., Porter, M. A., & Onnela, J.-P. (2010). Community Structure in Time-Dependent, Multiscale, and Multiplex Networks. *Science*, 328(5980), 876–878. <https://doi.org/10.1126/science.1184819>
- Nagahama, Y., Okina, T., Suzuki, N., & Matsuda, M. (2010). Neural correlates of psychotic symptoms in dementia with Lewy bodies. *Brain*, 133(2), 557–567. <https://doi.org/10.1093/brain/awp295>
- Nagahama, Y., Okina, T., Suzuki, N., Matsuda, M., Fukao, K., & Murai, T. (2007). Classification of Psychotic Symptoms in Dementia With Lewy Bodies. *The American Journal of Geriatric Psychiatry*, 15(11), 961–967. <https://doi.org/10.1097/JGP.0b013e3180cc1fdf>
- Nahab, F. B., Kundu, P., Gallea, C., Kakareka, J., Pursley, R., Pohida, T., Miletta, N., Friedman, J., & Hallett, M. (2011). The Neural Processes Underlying Self-Agency. *Cerebral Cortex*, 21(1), 48–55. <https://doi.org/10.1093/cercor/bhq059>
- Norman, K. A., Polyn, S. M., Detre, G. J., & Haxby, J. V. (2006). Beyond mind-reading: Multi-voxel pattern analysis of fMRI data. *Trends in Cognitive Sciences*, 10(9), 424–430. <https://doi.org/10.1016/j.tics.2006.07.005>
- Nuevo, R., Chatterji, S., Verdes, E., Naidoo, N., Arango, C., & Ayuso-Mateos, J. L. (2012). The Continuum of Psychotic Symptoms in the General Population: A Cross-national Study. *Schizophrenia Bulletin*, 38(3), 475–485. <https://doi.org/10.1093/schbul/sbq099>
- O'Callaghan, C., Hall, J. M., Tomassini, A., Muller, A. J., Walpola, I. C., Moustafa, A. A., Shine, J. M., & Lewis, S. J. G. (2017). Visual Hallucinations Are Characterized by Impaired Sensory Evidence Accumulation: Insights From Hierarchical Drift Diffusion Modeling in Parkinson's Disease.

References

- Biological Psychiatry: Cognitive Neuroscience and Neuroimaging*, 2(8), 680–688. <https://doi.org/10.1016/j.bpsc.2017.04.007>
- Ogawa, S., Lee, T. M., Kay, A. R., & Tank, D. W. (1990). Brain magnetic resonance imaging with contrast dependent on blood oxygenation. *Proceedings of the National Academy of Sciences*, 87(24), 9868–9872. <https://doi.org/10.1073/pnas.87.24.9868>
- Ogawa, S., & Lee, T.-M. (1990). Magnetic resonance imaging of blood vessels at high fields: In vivo and in vitro measurements and image simulation. *Magnetic Resonance in Medicine*, 16(1), 9–18. <https://doi.org/10.1002/mrm.1910160103>
- O'Reilly, J. X., Jbabdi, S., & Behrens, T. E. J. (2012). How can a Bayesian approach inform neuroscience?: How can a Bayesian approach inform neuroscience? *European Journal of Neuroscience*, 35(7), 1169–1179. <https://doi.org/10.1111/j.1460-9568.2012.08010.x>
- Orepic, P., Rognini, G., Kannape, O. A., Faivre, N., & Blanke, O. (2021). Sensorimotor conflicts induce somatic passivity and louden quiet voices in healthy listeners. *Schizophrenia Research*, 231, 170–177. <https://doi.org/10.1016/j.schres.2021.03.014>
- Orlov, N. D., Giampietro, V., O'Daly, O., Lam, S.-L., Barker, G. J., Rubia, K., McGuire, P., Shergill, S. S., & Allen, P. (2018). Real-time fMRI neurofeedback to down-regulate superior temporal gyrus activity in patients with schizophrenia and auditory hallucinations: A proof-of-concept study. *Translational Psychiatry*, 8(1), 46. <https://doi.org/10.1038/s41398-017-0067-5>
- Oros-Peusquens, A.-M., Loução, R., Abbas, Z., Gras, V., Zimmermann, M., & Shah, N. J. (2019). A Single-Scan, Rapid Whole-Brain Protocol for Quantitative Water Content Mapping With Neurobiological Implications. *Frontiers in Neurology*, 10, 1333. <https://doi.org/10.3389/fneur.2019.01333>
- Pagonabarraga, J., Martinez-Horta, S., Fernández de Bobadilla, R., Pérez, J., Ribosa-Nogué, R., Marín, J., Pascual-Sedano, B., García, C., Gironell, A., & Kulisevsky, J. (2016). Minor hallucinations occur in drug-naïve Parkinson's disease patients, even from the premotor phase: Minor Hallucinations in Untreated PD Patients. *Movement Disorders*, 31(1), 45–52. <https://doi.org/10.1002/mds.26432>
- Pagonabarraga, J., Soriano-Mas, C., Llebaria, G., López-Solà, M., Pujol, J., & Kulisevsky, J. (2014). Neural correlates of minor hallucinations in non-demented patients with Parkinson's disease. *Parkinsonism & Related Disorders*, 20(3), 290–296. <https://doi.org/10.1016/j.parkreldis.2013.11.017>
- Palaniyappan, L., & Liddle, P. F. (2012). Does the salience network play a cardinal role in psychosis? An emerging hypothesis of insular dysfunction. *Journal of Psychiatry and Neuroscience*, 37(1), 17–27. <https://doi.org/10.1503/jpn.100176>
- Palaniyappan, L., Simmonite, M., White, T. P., Liddle, E. B., & Liddle, P. F. (2013). Neural Primacy of the Salience Processing System in Schizophrenia. *Neuron*, 79(4), 814–828. <https://doi.org/10.1016/j.neuron.2013.06.027>
- Pamplona, G. S. P., Heldner, J., Langner, R., Koush, Y., Michels, L., Ionta, S., Scharnowski, F., & Salmon, C. E. G. (2020). Network-based fMRI-neurofeedback training of sustained attention. *NeuroImage*, 221, 117194. <https://doi.org/10.1016/j.neuroimage.2020.117194>

References

- Pankow, A., Katthagen, T., Diner, S., Deserno, L., Boehme, R., Kathmann, N., Gleich, T., Gaebler, M., Walter, H., Heinz, A., & Schlagenhauf, F. (2015). Aberrant Salience Is Related to Dysfunctional Self-Referential Processing in Psychosis. *Schizophrenia Bulletin*, sbv098. <https://doi.org/10.1093/schbul/sbv098>
- Park, H.-D., & Blanke, O. (2019). Coupling Inner and Outer Body for Self-Consciousness. *Trends in Cognitive Sciences*, 23(5), 377–388. <https://doi.org/10.1016/j.tics.2019.02.002>
- Peter Brugger, Marianne Regard, & Theodor Landis. (1996). Brugger et al 1996—Unilaterally felt presences The neuropsychiatry of ones invisible doppelganger.pdf. *Neuropsychiatry Neuropsychology and Behavioral Neurology*, 2, 19–38.
- Peter Brugger, Marianne Regard, Theodor Landis, & O Oelz. (1999). Brugger et al 1999—Hallucinatory experiences in extreme-altitude climbers.pdf. *Neuropsychiatry Neuropsychology and Behavioral Neurology*, 12, 67–71.
- Petkova, V. I., & Ehrsson, H. H. (2008). If I Were You: Perceptual Illusion of Body Swapping. *PLoS ONE*, 3(12), e3832. <https://doi.org/10.1371/journal.pone.0003832>
- Postuma, R. B., & Berg, D. (2016). Advances in markers of prodromal Parkinson disease. *Nature Reviews Neurology*, 12(11), 622–634. <https://doi.org/10.1038/nrneurol.2016.152>
- Potheegadoo, J., Dhanis, H., Horvath, J., Burkhard, P. R., & Blanke, O. (2022). Presence Hallucinations during Locomotion in Patients with Parkinson’s Disease. *Movement Disorders Clinical Practice*, 9(1), 127–129. <https://doi.org/10.1002/mdc3.13367>
- Press, C., Kok, P., & Yon, D. (2020). The Perceptual Prediction Paradox. *Trends in Cognitive Sciences*, 24(1), 13–24. <https://doi.org/10.1016/j.tics.2019.11.003>
- Preti, M. G., Bolton, T. A., & Van De Ville, D. (2017). The dynamic functional connectome: State-of-the-art and perspectives. *NeuroImage*, 160, 41–54. <https://doi.org/10.1016/j.neuroimage.2016.12.061>
- Raichle, M. E., MacLeod, A. M., Snyder, A. Z., Powers, W. J., Gusnard, D. A., & Shulman, G. L. (2001). A default mode of brain function. *Proceedings of the National Academy of Sciences*, 98(2), 676–682. <https://doi.org/10.1073/pnas.98.2.676>
- Ramirez-Ruiz, B., Junque, C., Marti, M.-J., Valldeoriola, F., & Tolosa, E. (2007). Cognitive Changes in Parkinson’s Disease Patients with Visual Hallucinations. *Dementia and Geriatric Cognitive Disorders*, 23(5), 281–288. <https://doi.org/10.1159/000100850>
- Ravina, B., Marder, K., Fernandez, H. H., Friedman, J. H., McDonald, W., Murphy, D., Aarsland, D., Babcock, D., Cummings, J., Endicott, J., Factor, S., Galpern, W., Lees, A., Marsh, L., Stacy, M., Gwinn-Hardy, K., Voon, V., & Goetz, C. (2007). Diagnostic criteria for psychosis in Parkinson’s disease: Report of an NINDS, NIMH work group. *Movement Disorders*, 22(8), 1061–1068. <https://doi.org/10.1002/mds.21382>
- Rosenberg, M. D., Finn, E. S., Scheinost, D., Papademetris, X., Shen, X., Constable, R. T., & Chun, M. M. (2016). A neuromarker of sustained attention from whole-brain functional connectivity. *Nature Neuroscience*, 19(1), 165–171. <https://doi.org/10.1038/nn.4179>

References

- Sadaghiani, S., Poline, J.-B., Kleinschmidt, A., & D'Esposito, M. (2015). Ongoing dynamics in large-scale functional connectivity predict perception. *Proceedings of the National Academy of Sciences*, 112(27), 8463–8468. <https://doi.org/10.1073/pnas.1420687112>
- Salomon, R., Progin, P., Griffo, A., Rognini, G., Do, K. Q., Conus, P., Marchesotti, S., Bernasconi, F., Hagmann, P., Serino, A., & Blanke, O. (2020). Sensorimotor Induction of Auditory Misattribution in Early Psychosis. *Schizophrenia Bulletin*, 46(4), 947–954. <https://doi.org/10.1093/schbul/sbz136>
- Schmack, K., Gomez-Carrillo de Castro, A., Rothkirch, M., Sekutowicz, M., Rossler, H., Haynes, J.-D., Heinz, A., Petrovic, P., & Sterzer, P. (2013). Delusions and the Role of Beliefs in Perceptual Inference. *Journal of Neuroscience*, 33(34), 13701–13712. <https://doi.org/10.1523/JNEUROSCI.1778-13.2013>
- Schnell, K., Heekeren, K., Daumann, J., Schnell, T., Schnitker, R., Möller-Hartmann, W., & Gouzoulis-Mayfrank, E. (2008). Correlation of passivity symptoms and dysfunctional visuomotor action monitoring in psychosis. *Brain*, 131(10), 2783–2797. <https://doi.org/10.1093/brain/awn184>
- Seeley, W. W., Menon, V., Schatzberg, A. F., Keller, J., Glover, G. H., Kenna, H., Reiss, A. L., & Greicius, M. D. (2007). Dissociable Intrinsic Connectivity Networks for Salience Processing and Executive Control. *Journal of Neuroscience*, 27(9), 2349–2356. <https://doi.org/10.1523/JNEUROSCI.5587-06.2007>
- Serino, A., Pozeg, P., Bernasconi, F., Solcà, M., Hara, M., Progin, P., Stripeikyte, G., Dhanis, H., Salomon, R., Bleuler, H., Rognini, G., & Blanke, O. (2021). Thought consciousness and source monitoring depend on robotically controlled sensorimotor conflicts and illusory states. *iScience*, 24(1), 101955. <https://doi.org/10.1016/j.isci.2020.101955>
- Shams, L., Kamitani, Y., & Shimojo, S. (2002). Visual illusion induced by sound. *Cognitive Brain Research*, 14(1), 147–152. [https://doi.org/10.1016/S0926-6410\(02\)00069-1](https://doi.org/10.1016/S0926-6410(02)00069-1)
- Sheng, J., Yan, Y., Yang, X., Yuan, T., & Cui, D. (2019). The effects of Mindfulness Meditation on hallucination and delusion in severe schizophrenia patients with more than 20 years' medical history. *CNS Neuroscience & Therapeutics*, 25(1), 147–150. <https://doi.org/10.1111/cns.13067>
- Shergill, S. S., Bays, P. M., Frith, C. D., & Wolpert, D. M. (2003). Two Eyes for an Eye: The Neuroscience of Force Escalation. *Science*, 301(5630), 187–187. <https://doi.org/10.1126/science.1085327>
- Shergill, S. S., Samson, G., Bays, P. M., Frith, C. D., & Wolpert, D. M. (2005). Evidence for Sensory Prediction Deficits in Schizophrenia. *American Journal of Psychiatry*, 162(12), 2384–2386. <https://doi.org/10.1176/appi.ajp.162.12.2384>
- Sherwood, M. S., Kane, J. H., Weisend, M. P., & Parker, J. G. (2016). Enhanced control of dorsolateral prefrontal cortex neurophysiology with real-time functional magnetic resonance imaging (rt-fMRI) neurofeedback training and working memory practice. *NeuroImage*, 124, 214–223. <https://doi.org/10.1016/j.neuroimage.2015.08.074>
- Shibata, K., Watanabe, T., Sasaki, Y., & Kawato, M. (2011). Perceptual Learning Incepted by Decoded fMRI Neurofeedback Without Stimulus Presentation. *Science*, 334(6061), 1413–1415. <https://doi.org/10.1126/science.1212003>

References

- Shine, J. M., Halliday, G. M., Gilat, M., Matar, E., Bolitho, S. J., Carlos, M., Naismith, S. L., & Lewis, S. J. G. (2014). The role of dysfunctional attentional control networks in visual misperceptions in Parkinson's disease: Visual Misperceptions in Parkinson's Disease. *Human Brain Mapping*, 35(5), 2206–2219. <https://doi.org/10.1002/hbm.22321>
- Shine, J. M., Halliday, G. M., Naismith, S. L., & Lewis, S. J. G. (2011). Visual misperceptions and hallucinations in Parkinson's disease: Dysfunction of attentional control networks?: Attentional Control Networks in PD Hallucinations. *Movement Disorders*, 26(12), 2154–2159. <https://doi.org/10.1002/mds.23896>
- Shine, J. M., Muller, A. J., O'Callaghan, C., Hornberger, M., Halliday, G. M., & Lewis, S. J. (2015). Abnormal connectivity between the default mode and the visual system underlies the manifestation of visual hallucinations in Parkinson's disease: A task-based fMRI study. *Npj Parkinson's Disease*, 1(1), 15003. <https://doi.org/10.1038/npjparkd.2015.3>
- Shirer, W. R., Ryali, S., Rykhlevskaia, E., Menon, V., & Greicius, M. D. (2012). Decoding Subject-Driven Cognitive States with Whole-Brain Connectivity Patterns. *Cerebral Cortex*, 22(1), 158–165. <https://doi.org/10.1093/cercor/bhr099>
- Sitaram, R., Ros, T., Stoeckel, L., Haller, S., Scharnowski, F., Lewis-Peacock, J., Weiskopf, N., Blefari, M. L., Rana, M., Oblak, E., Birbaumer, N., & Sulzer, J. (2017). Closed-loop brain training: The science of neurofeedback. *Nature Reviews Neuroscience*, 18(2), 86–100. <https://doi.org/10.1038/nrn.2016.164>
- Skudlarski, P., Jagannathan, K., Anderson, K., Stevens, M. C., Calhoun, V. D., Skudlarska, B. A., & Pearlson, G. (2010). Brain Connectivity Is Not Only Lower but Different in Schizophrenia: A Combined Anatomical and Functional Approach. *Biological Psychiatry*, 68(1), 61–69. <https://doi.org/10.1016/j.biopsych.2010.03.035>
- Smieskova, R., Roiser, J. P., Chaddock, C. A., Schmidt, A., Harrisberger, F., Bendfeldt, K., Simon, A., Walter, A., Fusar-Poli, P., McGuire, P. K., Lang, U. E., Riecher-Rössler, A., & Borgwardt, S. (2015). Modulation of motivational salience processing during the early stages of psychosis. *Schizophrenia Research*, 166(1–3), 17–23. <https://doi.org/10.1016/j.schres.2015.04.036>
- Sorger, B., Scharnowski, F., Linden, D. E. J., Hampson, M., & Young, K. D. (2019). Control freaks: Towards optimal selection of control conditions for fMRI neurofeedback studies. *NeuroImage*, 186, 256–265. <https://doi.org/10.1016/j.neuroimage.2018.11.004>
- Spetter, M. S., Malekshahi, R., Birbaumer, N., Lühns, M., van der Veer, A. H., Scheffler, K., Spuckti, S., Preissl, H., Veit, R., & Hallschmid, M. (2017). Volitional regulation of brain responses to food stimuli in overweight and obese subjects: A real-time fMRI feedback study. *Appetite*, 112, 188–195. <https://doi.org/10.1016/j.appet.2017.01.032>
- Squire, L. R. (2009). The Legacy of Patient H.M. for Neuroscience. *Neuron*, 61(1), 6–9. <https://doi.org/10.1016/j.neuron.2008.12.023>
- Sridharan, D., Levitin, D. J., & Menon, V. (2008). A critical role for the right fronto-insular cortex in switching between central-executive and default-mode networks. *Proceedings of the National Academy of Sciences*, 105(34), 12569–12574. <https://doi.org/10.1073/pnas.0800005105>

References

- Stefano Monti, Pablo Tamayo, Jill Mesirov, & Todd Golub. (2003). Consensus Clustering: A Resampling-Based Method for Class Discovery and Visualization of Gene Expression Microarray Data. *Machine Learning*, 52, 91–118.
- Stephan, K. E., Friston, K. J., & Frith, C. D. (2009). Dysconnection in Schizophrenia: From Abnormal Synaptic Plasticity to Failures of Self-monitoring. *Schizophrenia Bulletin*, 35(3), 509–527. <https://doi.org/10.1093/schbul/sbn176>
- Stoeckel, L. E., Garrison, K. A., Ghosh, S. S., Wightton, P., Hanlon, C. A., Gilman, J. M., Greer, S., Turk-Browne, N. B., deBettencourt, M. T., Scheinost, D., Craddock, C., Thompson, T., Calderon, V., Bauer, C. C., George, M., Breiter, H. C., Whitfield-Gabrieli, S., Gabrieli, J. D., LaConte, S. M., ... Evins, A. E. (2014). Optimizing real time fMRI neurofeedback for therapeutic discovery and development. *NeuroImage: Clinical*, 5, 245–255. <https://doi.org/10.1016/j.nicl.2014.07.002>
- Stripeikyte, G., Potheegadoo, J., Progin, P., Rognini, G., Blondiaux, E., Salomon, R., Griffa, A., Hagmann, P., Faivre, N., Do, K. Q., Conus, P., & Blanke, O. (2021). Fronto-Temporal Disconnection Within the Presence Hallucination Network in Psychotic Patients With Passivity Experiences. *Schizophrenia Bulletin*, 47(6), 1718–1728. <https://doi.org/10.1093/schbul/sbab031>
- Stuke, H., Weilhhammer, V. A., Sterzer, P., & Schmack, K. (2018). Delusion Proneness is Linked to a Reduced Usage of Prior Beliefs in Perceptual Decisions. *Schizophrenia Bulletin*. <https://doi.org/10.1093/schbul/sbx189>
- Subramanian, L., Hindle, J. V., Johnston, S., Roberts, M. V., Husain, M., Goebel, R., & Linden, D. (2011). Real-Time Functional Magnetic Resonance Imaging Neurofeedback for Treatment of Parkinson's Disease. *Journal of Neuroscience*, 31(45), 16309–16317. <https://doi.org/10.1523/JNEUROSCI.3498-11.2011>
- Sulzer, J., Haller, S., Scharnowski, F., Weiskopf, N., Birbaumer, N., Blefari, M. L., Bruehl, A. B., Cohen, L. G., deCharms, R. C., Gassert, R., Goebel, R., Herwig, U., LaConte, S., Linden, D., Luft, A., Seifritz, E., & Sitaram, R. (2013). Real-time fMRI neurofeedback: Progress and challenges. *NeuroImage*, 76, 386–399. <https://doi.org/10.1016/j.neuroimage.2013.03.033>
- Sulzer, J., Sitaram, R., Blefari, M. L., Kollias, S., Birbaumer, N., Stephan, K. E., Luft, A., & Gassert, R. (2013). Neurofeedback-mediated self-regulation of the dopaminergic midbrain. *NeuroImage*, 83, 817–825. <https://doi.org/10.1016/j.neuroimage.2013.05.115>
- Supekar, K., Cai, W., Krishnadas, R., Palaniyappan, L., & Menon, V. (2019). Dysregulated Brain Dynamics in a Triple-Network Saliency Model of Schizophrenia and Its Relation to Psychosis. *Biological Psychiatry*, 85(1), 60–69. <https://doi.org/10.1016/j.biopsych.2018.07.020>
- Tagliazucchi, E., Balenzuela, P., Fraiman, D., & Chialvo, D. R. (2012). Criticality in Large-Scale Brain fMRI Dynamics Unveiled by a Novel Point Process Analysis. *Frontiers in Physiology*, 3. <https://doi.org/10.3389/fphys.2012.00015>
- Tagliazucchi, E., Balenzuela, P., Fraiman, D., Montoya, P., & Chialvo, D. R. (2011). Spontaneous BOLD event triggered averages for estimating functional connectivity at resting state. *Neuroscience Letters*, 488(2), 158–163. <https://doi.org/10.1016/j.neulet.2010.11.020>

References

- Thibault, R. T., MacPherson, A., Lifshitz, M., Roth, R. R., & Raz, A. (2018). Neurofeedback with fMRI: A critical systematic review. *NeuroImage*, 172, 786–807. <https://doi.org/10.1016/j.neuroimage.2017.12.071>
- Thomas Yeo, B. T., Krienen, F. M., Sepulcre, J., Sabuncu, M. R., Lashkari, D., Hollinshead, M., Roffman, J. L., Smoller, J. W., Zöllei, L., Polimeni, J. R., Fischl, B., Liu, H., & Buckner, R. L. (2011). The organization of the human cerebral cortex estimated by intrinsic functional connectivity. *Journal of Neurophysiology*, 106(3), 1125–1165. <https://doi.org/10.1152/jn.00338.2011>
- Tracy, D., & Shergill, S. (2013). Mechanisms Underlying Auditory Hallucinations—Understanding Perception without Stimulus. *Brain Sciences*, 3(4), 642–669. <https://doi.org/10.3390/brainsci3020642>
- Uddin, L. Q. (2015). Salience processing and insular cortical function and dysfunction. *Nature Reviews Neuroscience*, 16(1), 55–61. <https://doi.org/10.1038/nrn3857>
- van Geuns, R.-J. M., Wielopolski, P. A., de Bruin, H. G., Rensing, B. J., Oudkerk, M., & de Feyter, P. J. (1999). Basic Principles of Magnetic Resonance Imaging. *Progress in CardioVascular Diseases*, 42(2), 149–156.
- van Wassenhove, V., Grant, K. W., & Poeppel, D. (2007). Temporal window of integration in auditory-visual speech perception. *Neuropsychologia*, 45(3), 598–607. <https://doi.org/10.1016/j.neuropsychologia.2006.01.001>
- Vercammen, A., de Haan, E. H. F., & Aleman, A. (2008). Hearing a voice in the noise: Auditory hallucinations and speech perception. *Psychological Medicine*, 38(8), 1177–1184. <https://doi.org/10.1017/S0033291707002437>
- Vidaurre, D., Smith, S. M., & Woolrich, M. W. (2017). Brain network dynamics are hierarchically organized in time. *Proceedings of the National Academy of Sciences*, 114(48), 12827–12832. <https://doi.org/10.1073/pnas.1705120114>
- Wackermann, J., Putz, P., & Allefeld, C. (2008). Ganzfeld-induced hallucinatory experience, its phenomenology and cerebral electrophysiology. *Cortex*, 44(10), 1364–1378. <https://doi.org/10.1016/j.cortex.2007.05.003>
- Watanabe, T., Sasaki, Y., Shibata, K., & Kawato, M. (2017). Advances in fMRI Real-Time Neurofeedback. *Trends in Cognitive Sciences*, 21(12), 997–1010. <https://doi.org/10.1016/j.tics.2017.09.010>
- Weiskopf, N. (2012). Real-time fMRI and its application to neurofeedback. *NeuroImage*, 62(2), 682–692. <https://doi.org/10.1016/j.neuroimage.2011.10.009>
- Weiskopf, N., Sitaram, R., Josephs, O., Veit, R., Scharnowski, F., Goebel, R., Birbaumer, N., Deichmann, R., & Mathiak, K. (2007). Real-time functional magnetic resonance imaging: Methods and applications. *Magnetic Resonance Imaging*, 25(6), 989–1003. <https://doi.org/10.1016/j.mri.2007.02.007>
- White, T. P., Joseph, V., Francis, S. T., & Liddle, P. F. (2010). Aberrant salience network (bilateral insula and anterior cingulate cortex) connectivity during information processing in schizophrenia. *Schizophrenia Research*, 123(2–3), 105–115. <https://doi.org/10.1016/j.schres.2010.07.020>

References

- Whitfield-Gabrieli, S., & Ford, J. M. (2012). Default Mode Network Activity and Connectivity in Psychopathology. *Annual Review of Clinical Psychology*, 8(1), 49–76. <https://doi.org/10.1146/annurev-clinpsy-032511-143049>
- Wolpert, D. M., Ghahramani, Z., & Jordan, M. I. (1995). An internal model for sensorimotor integration. *Science (New York, N.Y.)*, 269(5232), 1880–1882. <https://doi.org/10.1126/science.7569931>
- Wolpert, D. M., & Kawato, M. (1998). Multiple paired forward and inverse models for motor control. *Neural Networks*, 11(7–8), 1317–1329. [https://doi.org/10.1016/S0893-6080\(98\)00066-5](https://doi.org/10.1016/S0893-6080(98)00066-5)
- Wood, R. A., Hopkins, S. A., Moodley, K. K., & Chan, D. (2015). Fifty Percent Prevalence of Extracampine Hallucinations in Parkinson's Disease Patients. *Frontiers in Neurology*, 6. <https://doi.org/10.3389/fneur.2015.00263>
- Woodward, N. D., Rogers, B., & Heckers, S. (2011). Functional resting-state networks are differentially affected in schizophrenia. *Schizophrenia Research*, 130(1–3), 86–93. <https://doi.org/10.1016/j.schres.2011.03.010>
- Wurtz, R. H. (2009). Recounting the impact of Hubel and Wiesel: Recounting the impact of Hubel and Wiesel. *The Journal of Physiology*, 587(12), 2817–2823. <https://doi.org/10.1113/jphysiol.2009.170209>
- Yamada, T., Hashimoto, R., Yahata, N., Ichikawa, N., Yoshihara, Y., Okamoto, Y., Kato, N., Takahashi, H., & Kawato, M. (2017). Resting-State Functional Connectivity-Based Biomarkers and Functional MRI-Based Neurofeedback for Psychiatric Disorders: A Challenge for Developing Theranostic Biomarkers. *International Journal of Neuropsychopharmacology*, 20(10), 769–781. <https://doi.org/10.1093/ijnp/pyx059>
- Yamashita, A., Hayasaka, S., Kawato, M., & Imamizu, H. (2017). Connectivity Neurofeedback Training Can Differentially Change Functional Connectivity and Cognitive Performance. *Cerebral Cortex*, 27(10), 4960–4970. <https://doi.org/10.1093/cercor/bhx177>
- Yoo, S.-S., & Jolesz, F. A. (2002). Functional MRI for neurofeedback: Feasibility study on a hand motor task. *NeuroReport*, 13(11). https://journals.lww.com/neuroreport/Fulltext/2002/08070/Functional_MRI_for_neurofeedback_feasibility.5.aspx
- Yu, Q., Erhardt, E. B., Sui, J., Du, Y., He, H., Hjelm, D., Cetin, M. S., Rachakonda, S., Miller, R. L., Pearlson, G., & Calhoun, V. D. (2015). Assessing dynamic brain graphs of time-varying connectivity in fMRI data: Application to healthy controls and patients with schizophrenia. *NeuroImage*, 107, 345–355. <https://doi.org/10.1016/j.neuroimage.2014.12.020>
- Zarkali, A., Luppi, A. I., Stamatakis, E. A., Reeves, S., McColgan, P., Leyland, L.-A., Lees, A. J., & Weil, R. S. (2022). Changes in dynamic transitions between integrated and segregated states underlie visual hallucinations in Parkinson's disease. *Communications Biology*, 5(1), 928. <https://doi.org/10.1038/s42003-022-03903-x>
- Zilverstand, A., Sorger, B., Zimmermann, J., Kaas, A., & Goebel, R. (2014). Windowed Correlation: A Suitable Tool for Providing Dynamic fMRI-Based Functional Connectivity Neurofeedback on Task Difficulty. *PLoS ONE*, 9(1), e85929. <https://doi.org/10.1371/journal.pone.0085929>

References

- Zweerings, J., Hummel, B., Keller, M., Zvyagintsev, M., Schneider, F., Klasen, M., & Mathiak, K. (2019). Neurofeedback of core language network nodes modulates connectivity with the default-mode network: A double-blind fMRI neurofeedback study on auditory verbal hallucinations. *NeuroImage*, 189, 533–542. <https://doi.org/10.1016/j.neuroimage.2019.01.058>

Appendix

Appendix

6.1 TbCAPs: A toolbox for co-activation pattern analysis

Thomas A.W. Bolton ^{a,b,*}, Constantin Tuleasca ^{c,d,e,f}, Diana Wotruba ^g, Gwladys Rey ^h, Herberto Dhanis ⁱ, Baptiste Gauthier ⁱ, Farnaz Delavari ^j, Elenor Morgenroth ^a, Julian Gaviria ^h, Eva Blondiaux ⁱ, Lukasz Smigielski ⁱ, Dimitri Van De Ville ^{a, b}

^a Institute of Bioengineering, École Polytechnique Fédérale de Lausanne (EPFL), Lausanne, Switzerland

^b Department of Radiology and Medical Informatics, University of Geneva (UNIGE), Geneva, Switzerland

^c Centre Hospitalier Universitaire Bicêtre, Service de Neurochirurgie, Paris, France

^d Sorbonne Université, Faculté de Médecine, Paris, France

^e Lausanne University (UNIL), Faculty of Biology and Medicine (FBM), Lausanne, Switzerland

^f Centre Hospitalier Universitaire Vaudois (CHUV), Neurosurgery Service and Gamma Knife Center, Lausanne, Switzerland

^g Program for Sustainable Development of Mental Health, University of Zurich, Zurich, Switzerland

^h Department of Neuroscience, University of Geneva (UNIGE), Geneva, Switzerland

ⁱ Laboratory of Cognitive Neuroscience, École Polytechnique Fédérale de Lausanne (EPFL), Lausanne, Switzerland

^j Developmental Imaging and Psychopathology Laboratory, Office Médico-Pédagogique, Department of Psychiatry, University of Geneva (UNIGE), Geneva, Switzerland

* Corresponding author. Institute of Bioengineering, École Polytechnique Fédérale de Lausanne (EPFL), Lausanne, Switzerland.

E-mail address: thomas.bolton@epfl.ch (T.A.W. Bolton).

Appendix

Abstract

Functional magnetic resonance imaging provides rich spatio-temporal data of human brain activity during task and rest. Many recent efforts have focused on characterising dynamics of brain activity. One notable instance is co-activation pattern (CAP) analysis, a frame-wise analytical approach that disentangles the different functional brain networks interacting with a user-defined seed region. While promising applications in various clinical settings have been demonstrated, there is not yet any centralised, publicly accessible resource to facilitate the deployment of the technique.

Here, we release a working version of TbCAPs, a new toolbox for CAP analysis, which includes all steps of the analytical pipeline, introduces new methodological developments that build on already existing concepts, and enables a facilitated inspection of CAPs and resulting metrics of brain dynamics. The toolbox is available on a public academic repository at https://c4science.ch/source/CAP_Toolbox.git.

In addition, to illustrate the feasibility and usefulness of our pipeline, we describe an application to the study of human cognition. CAPs are constructed from resting-state fMRI using as seed the right dorsolateral prefrontal cortex, and, in a separate sample, we successfully predict a behavioural measure of continuous attentional performance from the metrics of CAP dynamics ($R = 0.59$).

Keywords: Dynamic functional connectivity, frame-wise analysis, co-activation pattern analysis, task-positive network, attention, continuous performance, open source software

Appendix

1.

Introduction

Functional magnetic resonance imaging (fMRI) has enabled to track temporal changes in activity levels at the whole-brain scale by means of the blood oxygenation level-dependent (BOLD) contrast, a proxy for neural activation (Logothetis et al., 2001). In addition to more traditional task-based studies in which BOLD changes are mapped to a paradigm of interest (Friston et al., 1994), the characterisation of statistical interdependence between remote brain locations—termed functional connectivity (Friston, 1994)—in the resting-state, and the concomitant definition of large-scale resting-state networks (RSNs), has been a popular endeavour (Biswal et al., 1995; FoX et al., 2005; Damoiseaux et al., 2006; Power et al., 2011), with great benefits for the understanding of cognition and disease (van den Heuvel and Hulshoff Pol, 2010; Greicius, 2008; FoX and Greicius, 2010).

Over the past years, it has become increasingly appreciated that cross- regional relationships do not remain static over the course of a full scanning session (Chang and Glover, 2010): instead, a given region rearranges its interactions along time, in ways that have been addressed with very diverse analytical tools (see Hutchison et al. (2013); Preti et al. (2017) for exhaustive reviews of the dynamic functional connectivity field). In one family of approaches that has been developed, it is assumed that only few salient time points contain the information of interest that shapes whole-brain correlational relationships; selecting only these frames, by means of a seed-based thresholding process, already enables to derive accurate RSN maps, even if as little as 10% of data points are retained (Tagliazucchi et al., 2012). The analysis then moves from a second-order correlation-based characterisation to a first-order activation viewpoint, and reduces computational load, a desirable feat in light of the numerous large-scale acquisition initiatives embraced by the fMRI community (Van Essen et al., 2013; Nooner et al., 2012; Holmes et al., 2015).

Building on this point process analysis concept, and inspired by the dynamic viewpoint on resting-state brain function, Liu and Duyn (2013) hypothesised that at different moments in time, the seed region of interest would display distinct interactions with the rest of the brain. A k-means clustering step was thus appended to frame selection, so that fMRI volumes with a large enough seed activity would be partitioned into a limited set of co-activation patterns (CAPs).

Since then, co-activation pattern analysis has started to gain momentum as a potent tool to reveal functional brain dynamics subtleties: analyses taking the posterior cingulate cortex (PCC) as a seed revealed alterations of spatial intensity level and occurrence in specific CAPs (Amico et al., 2014; Di Perri et al., 2018), while in adolescent depression, Kaiser et al. (2019) showed that the time spent in a specific frontoinsula-default network CAP positively correlated with symptoms severity. In other work, the renormalization of CAP occurrences in patients with essential tremor following surgical intervention could be tracked (Tuleasca et al., 2019).

In parallel to clinical applications, the technical details of the approach have also been addressed, in terms of retaining activation versus deactivation time points (Di and Biswal, 2015), extending it to the wholebrain (Liu et al., 2013), designing novel metrics of interest (Chen et al., 2015), or constraining the extent of spatial overlap across CAPs (Zhuang et al. 2018). For more details, the reader is pointed at the recent review of Liu et al. (2018).

Here, we wish to further foster the development of CAP analysis by releasing a dedicated toolbox, which enables to easily navigate through the steps of the analytical pipeline through a graphical user interface, and also offers additional technical developments regarding frame selection and metrics

computation. While the mathematical underpinnings of CAP analysis are relatively straightforward, we hope that providing such a resource will encourage practitioners to embrace the method, and that it will become easier to compare CAP analyses based on subtle, but sometimes important, differences in the processing pipeline. Through this resource, we also aim at preventing the variability in analytical results that may otherwise arise due to implementation differences alone (Bowring et al., 2019).

In addition, to exemplify the use of our toolbox, we describe an application of CAP analysis in the yet unaddressed setting of predicting cognitive skills: in a battery of healthy individuals, we show that continuous performance in a visual attention and vigilance task correlates with the expression profile of task-positive network (TPN) CAPs.

2. Material and methods

2.1 Co-activation pattern analysis theory

Let us consider the data matrix $X_S \in R^{V \times T}$, for subject s , where V is the number of voxels to consider in the analysis and T the number of time points. Each voxel-wise time course is temporally z-scored,

so that $\mu_i = \frac{\sum_{t=1}^T X_S(i,t)}{T}$ and $\sigma_i = \sqrt{\frac{\sum_{t=1}^T (X_S(i,t) - \mu_i)^2}{T-1}} = 1$, for all $i = 1, 2, \dots, V$.

Co-activation pattern analysis requires the definition of a seed region, whose interactions with the rest of the brain will be probed. Formally, a set of voxels \mathcal{L} that one wishes to consider is specified, and a time point t of the seed activation time course is then given by:

$$S_{seed}(t) = \frac{\sum_{i \in \mathcal{L}} X_S(i, t)}{|\mathcal{L}|}, \text{ for all } t \in 1, 2, \dots, T$$

Only time points when the seed time course takes sufficiently extreme values (denoting significant seed (de)activation) are considered. Let the activation threshold be T , we then construct the subject-specific set of time points \mathbb{T}_s that satisfies $S_{seed}(t) > T$ (if we wish to consider solely activation moments) or $S_{seed}(t) < -T$ (if we are interested in seed deactivation time points).

In this work, in addition to the above standard CAP methodology, we propose an extension in which more than one seed region can be considered: for each seed j and subject s , a set of time points $\mathbb{T}_{s,j}$ is derived. Assuming J separate seeds, one can then consider the time points when all seed time courses jointly take extreme values:

$$\mathbb{T}_{s,intersection} = \bigcap_{j=1}^J \mathbb{T}_{s,j}$$

Alternatively, one may instead be interested in the moments when at least one of the seed regions becomes strongly (de)active:

$$\mathbb{T}_{s,intersection} = \bigcup_{j=1}^J \mathbb{T}_{s,j}$$

Appendix

Finally, other additional criteria can be incorporated at the time point selection step: for instance, given the deleterious impact that head motion exerts on BOLD signals even following standard preprocessing (Power et al., 2012; Van Dijk et al., 2012; Satterthwaite et al., 2012), it may be desirable to only retain the frames for which framewise displacement does not exceed a threshold M .

After having selected the frames to keep for each subject, the next step is the population-level clustering of data points into CAPs. K-means clustering is used for this purpose, to optimise:

$$\underset{C}{\operatorname{argmin}} \sum_{k=1}^K \sum_{s=1}^S \sum_{t \in \mathcal{T}_S \cap \mathcal{L}_k} \operatorname{dist}(X_s(\cdot, t), c_k)$$

where K is the number of co-activation patterns to derive, $C = \{\mathcal{L}_1, \dots, \mathcal{L}_K\}$ summarizes the hard assignment of the frames to each CAP, and c_k is the spatial map for co-activation pattern k . The dist function depends on the type of distance to use in the algorithm. In addition, since k-means clustering is an iterative process with no guaranteed convergence towards the global optimum, the algorithm is run n_{rep} times.

In several previous works using CAPs, it was also suggested to solve Equation (4) after setting to 0 the voxel intensity values that, for each frame of interest, would not be part of the largest P_P or P_N percents—for positive-valued and negative-valued voxels, respectively (Liu and Duyn, 2013; Liu et al., 2013). Table 1 summarises the different parameters that are defined for CAP analysis, and also highlights the default values that we used in this work.

2.2 Metrics characterizing CAP dynamics

Once all retained frames have been assigned to CAP representatives, it becomes possible to construct, for each subject, an empirical transition probability matrix A_s that summarises the likelihood to transit from a given CAP at time t to another at time $t + 1$. Another available piece of information regards the likelihood to transit from and back to the baseline state (when the seed was not significantly (de)active). Further, if separate subject populations are used in computing CAPs and deriving associated metrics (as in our example application below), there are also occurrences of entries into an extra state associated to frames that could not be matched to any CAP with sufficient certainty.

An indicative example of averaged transition probability matrix across subjects is displayed in Fig. 1A (left column). Individual elements of the transition probability matrix may be considered as such (Chen et al., 2015), which would amount to a total of K^2 values per subject. To meaningfully lower the amount of features of interest, we propose to rather view the available information as a directional graph representation, from which a series of summarising metrics can be derived (Rubinov and Sporns, 2010). First, by sampling the diagonal elements of the matrix, we obtain a measure of resilience for each CAP: that is, the likelihood to remain in the same configuration from time t to $t + 1$. Second, after having set the diagonal elements of the matrix to 0, we can define the in-degree k_{in} (how likely a CAP is visited from any other), the out-degree k_{out} (how likely a CAP is exited towards any other), and the betweenness centrality—how important a CAP is regarding the shortest paths

Appendix

between other pairs (Freeman, 1979). In total, the feature space has thus been reduced from K^2 to 4 .K. This alternative viewpoint is exemplified in Fig. 1A (right column).

Parameter	Description	Default value
J	Number of seeds to use	1
\mathcal{L}	Voxel set to use as seed	Right dorsolateral pre frontal cortex
Polarity	Sign of the seed excursions to consider	Activation
Seed combination	Whether all or at least on see should be (de)active to retain a time point	n.a.
T	Threshold for frame selection	1,5
M	Threshold of framewise displacement	0.3 mm
K	Number of clusters to use	16
n_{rep}	Number of replicates of k-means	50
P_p	Percentage of positive valued voxels to keep in each frame for clustering	100
P_N	Percentage of negative valued voxels to keep in each frame for clustering	100
dist	Distance measure used for corr clustering	

In several works, counts or occurrences (that is, how many times a given CAP is expressed) were used as metrics of interest (Di Perri et al., 2018; Kaiser et al., 2019; Tuleasca et al., 2019). We verified that our suggested metrics also include the information rendered by the counts: as seen in Fig. 1B, the average correlation across CAPs between counts and in-degree, out-degree or resilience exceeded 0.8 (respectively $\rho = 0.83 \pm 0.11$, $\rho = 0.85 \pm 0.08$ and $\rho = 0.81 \pm 0.07$). From pair-wise comparisons between our four metrics, it can also be seen that in-degree and out-degree are strongly correlated ($\rho = 0.87 \pm 0.1$), while resilience and betweenness centrality capture separate information given their more moderate correlations (for resilience: $\rho = 0.45 \pm 0.11$, $\rho = 0.5 \pm 0.11$ and $\rho = 0.3 \pm 0.19$ with in-degree, out-degree and betweenness centrality, respectively; for betweenness centrality: $\rho = 0.59 \pm 0.13$ and $\rho = 0.59 \pm 0.13$ with in-degree and out-degree, respectively). Despite their overall similarity, we

Appendix

decided to retain both in-degree and out-degree as they still yielded different values in specific CAP cases.

In addition to the above metrics that summarise the transitory behaviour across different CAPs, an interesting complement is the assessment of which CAPs are entered from the baseline state of seed activity, as well as of which CAPs are the ones expressed just before a return to baseline activity. With this additional information, a total of $6 \cdot K$ features of interest are available per subject (as 6 metrics are computed for each of K different CAPs). These are the summarising measures that we use in our example application.

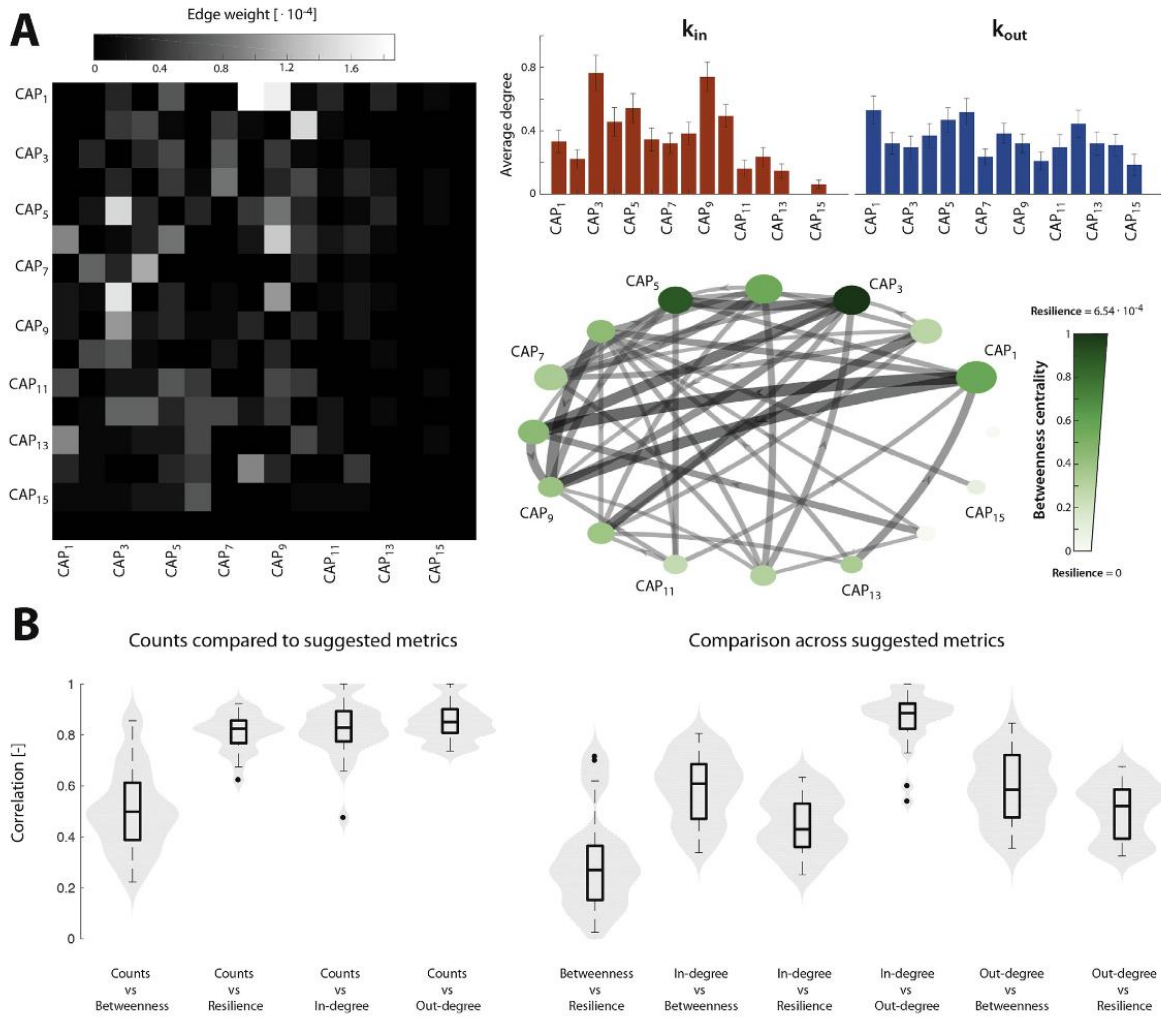


Fig. 1. Generation of CAP metrics. (A) Transitions across CAPs can be viewed in terms of individual transition probabilities for a total of K^2 features (left), or alternatively, a directional graph representation can be constructed (right) to extract in-degree (red bars), out-degree (blue bars), betweenness centrality (green colour coding of the nodes in the bottom plot) and resilience (size of the nodes) information, for a reduced total of $4 \cdot K$ features. In this example, which considers $K = 16$ CAPs, the feature space would thus be lowered from 256 to 64 (four-fold reduction). Error bars denote standard error of the mean, and the displayed transition probabilities, betweenness centrality values and resilience values are averages across subjects. (B) The extracted metrics contain equivalent information compared to the more traditional occurrences, and each such metric characterises partly different aspects of CAP dynamics, as seen by moderate correlations. Each box plot/violin plot representation depicts correlation values across $K = 16$ CAPs. Illustrations from this figure are generated from the data presented in our example application, with $A_P = 0\%$ (i.e., very lenient frame assignment), and without assigning scrubbed volumes.

2.3 *TbCAPs: implementation*

We implemented the CAPs processing pipeline as a toolbox in Matlab version 2017a (The MathWorks, Natick, USA). This software is freely accessible at https://c4science.ch/source/CAP_Toolbox.git. It contains a graphical user interface to facilitate the different steps of the pipeline. In addition, we also provide a scripted version of a typical analysis pipeline for power-users. An illustrative display of the graphical user interface at the end of a typical analysis is provided in Fig. 2. Next, we concisely describe the steps to be performed by the user, and the available options at each stage of the analysis. For more details on all existing functionalities, alternative example applications of the toolbox in clinical settings, and more specific suggestions regarding data preparation and quality control based on our past experience, the reader is pointed at the TbCAPs User Manual that accompanies this work as Supplementary Material.

2.3.1 *Data loading*

Prior to CAP analysis, the data at hand should have already underwent standard resting-state fMRI preprocessing steps, such as realignment, co-registration, regression of covariates of no interest, and filtering (we advise to only high-pass filter the data). All the volumes to analyse should have been warped to MNI space (as co-activation patterns will be derived from the whole population data). Particular care should be taken, during preprocessing, to attenuate physiology-related artefacts as much as possible, as they may otherwise exert pervasive effects on the BOLD signal (see Caballero-Gaudes and Reynolds (2017) for a recent review). Selected example strategies include the application of independent component analysis-based denoising approaches (Griffanti et al., 2014; Pruim et al., 2015), or the use of a set of regressors reflecting physiological variables (Glover et al., 2000).

Before loading the data to analyse in the toolbox, the user should first define how to mask it (that is, which voxels should be part of the analysis, excluding for instance out-of-brain ones); to do so, a popup window enables to choose between several mask options, after which the A1. Set mask button should be pressed. The user is then prompted to select any directory containing part of the functional data to analyse: this will enable to convert the chosen mask into the resolution of the functional data. The prefix specifying the data of interest (e.g., “sw” for an SPM preprocessing with warping and smoothing) should be provided through a dedicated text box.

In a second step, clicking the A2. Load data button prompts the user to select all the directories containing the functional data to analyse as part of a given group. We assume here that the data is arranged in a BIDS format (Gorgolewski et al., 2016), which implies in particular that different series of functional volumes to analyse should be located in different directories. Following loading, the data is z-scored by the toolbox, making it fully ready for CAP analysis. In addition, a text file summarising the results from the realignment step (that is, containing the 6 motion time courses) should be present in each directory, so that a framewise displacement time course can be constructed and enable subsequent scrubbing of corrupted frames; if such a file is not available, null motion is assumed and the analysis continues nonetheless.

2.3.2 Spatio-temporal selection

Following data loading, the user is prompted to select one or more seeds to use in the analysis (B. Select seed file(s) button): all seed files should be entered at once, each as an MNI space NIFTI volume, which does not need to be at the same spatial resolution as the functional data (this is automatically handled by the toolbox). The brain areas covered by the seed(s) can be inspected in brain slice representations, which can be navigated through by means of dedicated sliders. For now, we allow up to three separate seeds to be entered for the analyses. In addition, the interested user can also plot the average seed-based correlation map across subjects associated to the first selected seed.

The next step is to select which types of events should be retained (activation versus deactivation), and if more than one seed was selected, whether all seeds should show an extreme event at once to select a time point (Intersection option), or if a frame should be kept as long as at least one does so (Union option). The user is also prompted to determine the threshold T to use in selecting frames (or alternatively, a percentage P of most (de)active frames to retain), and the threshold M above which frames will be deemed excessively corrupted by head motion, and scrubbed out.

At the end of this process, clicking on the Select time points button performs the frame selection process, and summarises the percentage of kept volumes across subjects in a violin plot representation. Note that in lieu of a seed-based analysis, we also implemented an alternative seed-free option, following Liu et al. (2013), where frames are retained regardless of any seed. To run this option, the Seed-free analysis button should be clicked instead of seed loading.

2.3.3. Generation of co-activation patterns

Regarding the subsequent generation of CAPs, if the optimal number of clusters K to select is not known a priori, we offer the possibility to run consensus clustering (Monti et al., 2003), where clustering is run many times from $K = 2$ to a user-specified K_{\max} using a subsample of the data (the percentage of data points to use is specified by P_{CC} , and the number of iterations by N). A good clustering solution is one for which across folds, two frames are either always clustered together, or never clustered together (but not an intermediate case). We quantify this by the percentage of Ambiguously Clustered pairs, or PAC (Senbabaoglu et al., 2014), and display the stability measure $1 - PAC$. EXTENDED details on consensus clustering can be found in the TbCAPs User Manual.

Following the definition of how many CAPs should be extracted (parameter K in the interface), k-means clustering can be run by clicking the Cluster button. The first 5 CAPs with most occurrences across the subject population are displayed, and can be visually inspected and navigated through by dedicated sliders. As a complement, the matrix of spatial similarity between CAPs is also provided, and if using the Union option in a multi-seed analysis, the user is shown pie charts summarising, for each CAP, what fraction of frames was selected in a given seed combination configuration.

2.3.4. Computation of metrics

Finally, upon clicking the Compute metrics button, displays of CAP expression time courses and cumulative CAP expression along time appear on screen. The latter can be adjusted to selectively view

Appendix

the cumulative counts of a given CAP, across subjects and as a population average. Transition probability matrices can also be inspected in terms of average entries at the population level, or for each subject, with the option to select the one to display the data for.

In addition, 6 violin plot representations summarising the distribution of computed metrics across subjects for each CAP are also provided: they reflect (1) raw counts, (2) number of entries in a given CAP, (3) resilience, (4) betweenness centrality, (5) in-degree and (6) out-degree.

2.3.5. Analysis of multiple subject populations

In some settings, the user may wish to compare different subject populations (for instance, healthy controls and a clinical group): this can be done by sequentially loading up to 4 different populations at the start of the analysis. CAPs will be derived from the first population, and there is then the need to assign the frames from the other populations to the CAPs by a matching process. In doing so, the spatial correlation between a frame to assign and the CAP to which it is most similar is compared to the distribution of spatial correlations of the frames from population 1 that belong to the CAP in question: if the A_p^{th} percentile of this distribution is exceeded, assignment is performed; else, the frame is left unassigned and belongs to an extra $(K+1)^{\text{th}}$ cluster. Additional details regarding multi-population analyses are provided in the TbCAPs User Manual.

2.4. Application to experimental fMRI data

2.4.1. Functional data preprocessing

As a proof of feasibility and application of TbCAPs, we considered a sample of 181 subjects from the Human Connectome Project (Van Essen et al., 2013), aged between 26 and 35 years old. This sample originated from a slightly larger, randomly selected set of subjects that had at least one fully exploitable resting-state scanning session on which to apply the method, and less than 5% of recorded behavioral entries that were missing; a few subjects from this original set were discarded due to errors in preprocessing. Details regarding acquisition parameters can be found elsewhere (Smith et al., 2013), but briefly, the data was acquired at a TR of 0.72 s over 15 min (for a total of 1200 fMRI volumes), with a spatial resolution following initial preprocessing steps of 2 mm x 2 mm x 2 mm. We started from the minimally preprocessed resting-state data (first session, LR acquisition direction). The first 10 samples of the data were discarded. We then performed linear detrending, and regressed out low-frequency components of the discrete cosine transform basis with a cutoff frequency at 0.01 Hz. Due to collinearity with this basis, we did not regress out average white matter or cerebrospinal fluid time courses. We also chose not to regress motion parameter time courses, as motion is handled within the co-activation pattern pipeline by scrubbing, and because recent evidence points at the fact that motion regression schemes may not always be beneficial in the context of brain/behaviour analyses (Bolton et al., 2020). As for global signal regression, given the lack of a clear consensus (Murphy and Fox, 2017), we preferred to leave the data as untouched as possible and did not include it.

Following the regression step, the data was scrubbed at a framewise displacement threshold of 0.3 mm, and excised volumes were estimated with cubic spline interpolation. Although scrubbing is

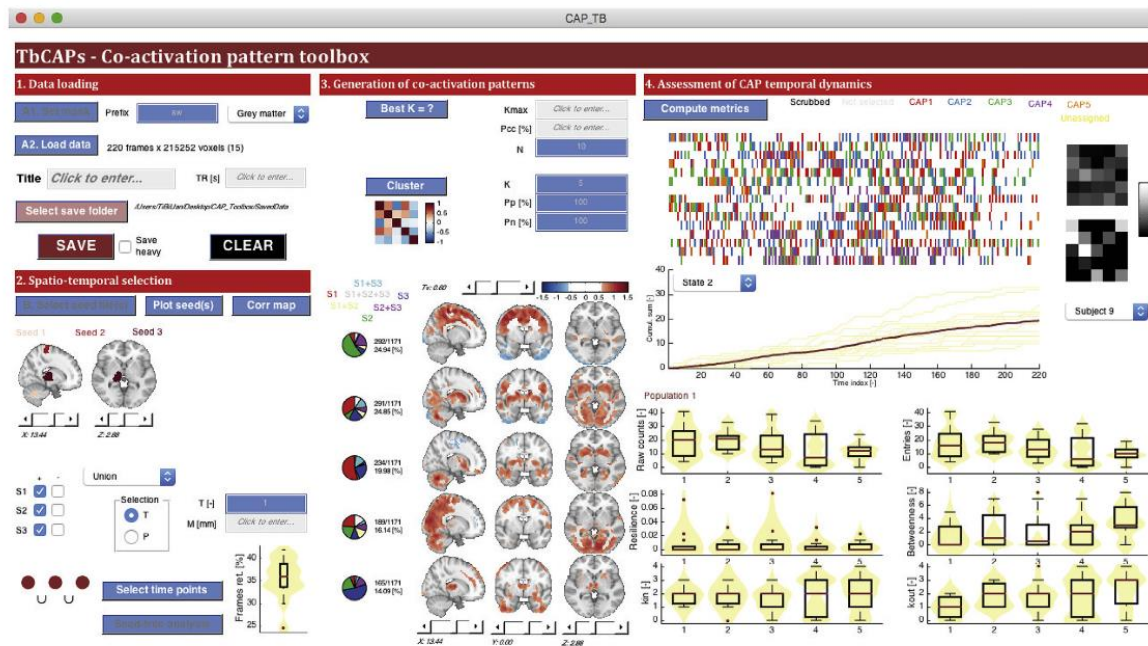


Fig. 2. Illustration of the TbCAPs graphical user interface. The typical output of a CAP analysis is displayed for a set of healthy volunteers, with three seeds chosen within a functional circuitry associated to essential tremor. This data was explicitly analysed in [Tuleasca et al. \(2019\)](#).

performed within TbCAPs, we reasoned that if we wished to try and assign scrubbed frames to CAPs in our additional analyses regarding head motion, it would make more sense to have previously corrected these volumes to the best of our abilities.

Then, individual fMRI volumes were smoothed at a full width at half maximum value of 5 mm, and in order to make the analyses computationally more affordable, spatial resolution was downsampled at 3 mm × 3 mm x 3 mm. Eventually, z-scoring was handled within TbCAPs as a final preprocessing step prior to CAP analysis.

2.4.2. Selection of seed and behaviour of interest

As a behaviour of interest to study, we selected the Short Penn Continuous Performance Test (SCPT), which quantifies continuous sustained attention ([Gur et al., 2010](#)). In more details, participants see vertical and horizontal red lines flash on screen, and from block to block, must respond either when the lines form a number, or a letter. The lines are displayed for 300 ms, followed by a 700 ms inter-trial interval.

We started from raw behavioral entries provided by the HCP, for 951 different subjects. There are 8 available SCPT measures: amount of true positives, false positives, true negatives or false negatives, median response time for true positive responses, sensitivity, specificity and longest run of non-responses. In order to reduce this information into one summary measure while filling in missing behavioural entries, we performed probabilistic PCA ([Bishop, 1999](#)). The output composite score positively correlated with true positives, true negatives, sensitivity and specificity ($\rho = 0.24, 1.00, 0.25$ and 1.00 , respectively), thus summarising overall task performance. We z-scored this output measure across subjects, in order to quantify performance with respect to the overall population. We then extracted the behavioural data related to the 181 subjects considered in this work.

Appendix

To study sustained attention, we focused on a right dorsolateral prefrontal cortex seed from the task-positive network, which we extracted from the associated independent component map provided by Shirer et al. (2012). Our hypothesis was that the expression of different TPN configurations would relate to sustained attentional performance.

2.4.3. Co-activation pattern analysis details

We resorted to a threshold $T = 1.5$ to select active frames, and performed scrubbing with a framewise displacement threshold $M = 0.3$ mm. To avoid double dipping (Kriegeskorte et al., 2009), CAPs were extracted from a randomly selected subset of 100 subjects, while we performed correlations with behaviour for the remaining 81 only. To determine the optimal number of clusters, we used consensus clustering (Monti et al., 2003). We then ran k-means $n_{rep} \propto 50$ separate times, keeping the best solution. We included all voxels in the analyses ($P_p = P_N = 100\%$), and used spatial correlation as our distance measure; given two similarly-sized vectors x and y , this thus yields $\text{dist}(x; y) = 1 - \text{corr}(x; y)$.

Following the extraction of the CAPs on our 100 training subjects, we determined which CAP was expressed at each retained fMRI volume of the other 81 subjects. To do so, we used the forementioned assignment process with AP ranging from 0 to 100%.

2.4.4. Assessment of brain/behaviour relationship

As imaging metrics of interest, we considered in-degree, out-degree, betweenness centrality and resilience for each CAP, and also included the amount of excursions from the baseline state, and the amount of excursions back to the baseline state. Thus, we generated a total of 6 . K imaging features per subject.

After having obtained the behavioral scores $\mathbf{b} \in \mathbb{R}^{81 \times 1}$ and metrics $\mathbf{M} \in \mathbb{R}^{81 \times 1 \times k}$ for our population of subjects, we used Partial Least Squares (PLS) analysis (McIntosh and Lobaugh, 2004; Krishnan et al., 2011) to probe the existence of a brain/behaviour relationship.

Briefly, consider a matrix of behavioural features $\mathbf{b} \in \mathbb{R}^{S \times n_B}$ and a matrix of imaging metrics $\mathbf{M} \in \mathbb{R}^{S \times n_M}$. Assuming that $n_B < n_M$, and using the singular value decomposition, the covariance between these two sets is given by:

$$R = \mathbf{M}^T \mathbf{B} = \mathbf{U} \Sigma \mathbf{V}^T = \sum_{i=1}^{n_B} \sigma_i \mathbf{u}_i \mathbf{v}_i^T$$

where each column in \mathbf{U} and \mathbf{V} contains the weights (so called saliences) that respectively multiply imaging and behavioural markers to yield a maximised covariance between both sets. The associated singular value σ_i is proportional to the fraction of covariance explained by the component at hand.

In our case, since $n_B = 1$ (we only consider one behavioural measure), only one covariance component is retrieved, which implies $v_1 = 1$. The interesting information lies in \mathbf{u}_1 : positive-valued saliences

Appendix

highlight metrics that are larger in subjects who show a greater cognitive ability, and negative-valued saliences are associated to metrics that, when larger, impede attentional performance.

Prior to running the algorithm, each of the 6 K features was z-scored across subjects. In order to assess significance, we reran PLS 1000 times after having randomly shuffled the subject entries in one of the two matrices; to non-parametrically derive a p-value, the singular value of the actual covariance component was compared to the null distribution constructed from this permutation process.

Further, to establish the stability of the salience weights, we reran PLS 1000 times using a randomly selected subsample of 80% of the data, and computed a bootstrap score for each salience weight as its mean across folds divided by its standard deviation.

2.4.5. Influence of head motion on quantified metrics

While scrubbing enables to minimise the deleterious impacts of motion on the analysis and compute clean CAPs, discarding frames also has the potential to distort transition probability estimates. For example, a succession of three frames in the same state (which would amount to a higher resilience for the CAP in question) would not be captured if the middle frame is scrubbed out.

To verify that our findings were minimally sensitive to this effect, we ran another series of analyses in which we also performed the afore- mentioned assignment process on scrubbed frames, with a similar AP range as for assigning test subject frames. This way, frames strongly distorted by head motion still do not enter the analysis, but more mildly affected fMRI volumes can be matched to their CAP. We verified the reproducibility of our findings upon this additional analytical step.

3. Results

Consensus clustering results are displayed in Fig. 3A for K values ranging from 10 to 40. Positive peaks highlight good candidate values (see figure legend for details); such values are present for diverse numbers of clusters (more notably at K 16; 22; 32). While a lower number of clusters yields a reduced feature space and more interpretable outcomes, CAPs may not be segmented finely enough to resolve insightful dynamic properties regarding cognition. Our strategy was thus to first perform an exploratory assessment, in which we evaluated the significance of the brain/behaviour correlation across a set of candidate K values ($K_{\text{opt}} = \{14; 16; 22; 32\}$) and assignment thresholds (forcing the assignment of all frames, or using $T_p = [0 : 5 : 100]$), to select the relevant parameters to proceed forward with.

The results of this exploratory assessment are displayed in Fig. 3B, when scrubbed frames are discarded (left panel) or also assigned to the CAPs at threshold AP (right panel). Both settings yield very similar significance values, which is good evidence that remaining head motion effects only have a minor influence on the analyses. As the assignment threshold increases (that is, less and less frames are assigned because the criterion becomes more and more stringent), significance generally decreases. A smooth spot can be observed for $K = \{16; 22\}$ and $A_p < 15\%$, which indicates that this granularity is optimal in the context of behavioural prediction. We selected $K = 16$ and $A_p 0\%$ as values for more detailed subsequent analyses.

Appendix

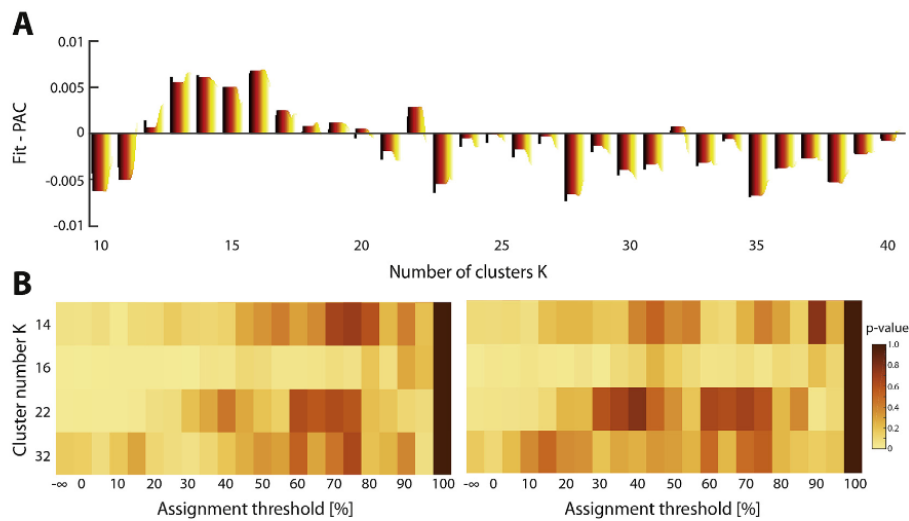


Fig. 3. Parameter selection. (A) Consensus clustering results for a range of K values from 10 to 40. Outputs from the algorithm are percentages of ambiguously clustered pairs (PAC); since this measure requires the definition of an interval of consensus values within which clustering is deemed *ambiguous*, we investigated a range of values as colour-coded from black to yellow. For each value, we fitted a decreasing exponential to capture the overall tendency of PAC values going down with larger cluster number. The y-axis of the plot depicts the difference between this fitted value and the actual value: thus, a positive difference means that the considered cluster number is more satisfying than what would be predicted in terms of the overall behaviour. (B) Across candidate cluster number values selected from consensus clustering, and assignment threshold values, significance of the relationship between CAP metrics and attentional abilities, as quantified by the p-value obtained upon PLS analysis. The left panel depicts the results for which scrubbed frames are not considered at all, while in the right panel, scrubbed frames were also assigned to the CAPs. The infinity symbol is used to depict a case in which assignment is done for all frames, even if a frame is not sufficiently close to any CAP when comparing its spatial correlation to the associated correlation distribution.

CAPs are displayed, for this chosen parameter set, in Fig. 4A, while their involvement in driving the brain/behaviour relationship, as quantified by salience weights across our range of investigated metrics, is depicted in Fig. 4B. The correlation between actual and predicted attentional performance was strongly significant ($R = 0.587$, $p < 0.001$; Fig. 4C). The associated covariance component found by PLS analysis was significant at $p = 0.003$. Note that this relationship is derived from only subjects that were not used to construct the CAPs.

CAP₁ depicts co-activation of a range of resting-state networks, including the auditory, somatomotor, visual and salience ones. Attentional performance was better in the subjects that transitioned more frequently from the baseline state of seed activity to this CAP. CAP₁ was also more often the entry point towards other CAPs in high performance subjects, as seen from a strongly positive out-degree salience weight.

Good subjects in terms of continuous performance also more often entered CAP₂ and CAP₇ from other states (as seen from positive in- degree salience weights), and these same 2 CAPs were also more influential in the transitory behaviour of CAP dynamics (since betweenness centrality salience weights also showed large positive values). In both CAPs, the seed region co-activates with a restricted set of areas including the right inferior parietal cortex (for both), the posterior cingulate cortex and medial prefrontal cortex (for CAP₂), and the right anterior prefrontal cortex (for CAP₇).

CAP₃ and CAP₅ were associated to good attentional abilities from the viewpoint of several metrics, which emphasises the importance of their expression: for CAP₃, it involved resilience, in-degree and out-degree, while for CAP₅, it was return to baseline, resilience, in-degree and betweenness centrality. Both CAPs include strong co-activation with the right inferior parietal cortex, and for CAP₃, also with the left cerebellum lobule VI and a subpart of the occipital cortex.

CAP₄ and CAP₁₄ were the only states whose expression was detrimental to attentional performance, in terms of betweenness centrality for the former, and of return to baseline for the latter. CAP₄

displayed bilateral right superior cortex and anterior prefrontal cortex co-activation with the seed, while for CAP₁₄, involved areas were the fusiform gyrus, parahippocampal cortex, and a diffuse right lateralized spot covering parts of the auditory, secondary somatosensory and posterior insular cortices.

The majority of the other CAPs that did not show any link to attentional performance involved co-activations with regions that were not part of the attentional networks: for instance, CAP₆ includes the anterior cingulate cortex and anterior insula; CAP₈ contains the anterior cingulate, visual and right somatosensory cortices; CAP₉ showcases primary visual and auditory cortices; CAP₁₀ shows the angular gyrus and part of the precuneus; and CAP₁₁ and CAP₁₅ mostly highlight ventral medial prefrontal cortex signal.

4. Discussion

In this work, we have introduced TbCAPs, a toolbox for co-activation pattern analysis, which provides practitioners with an intuitive dedicated graphical user interface as well as a powerful scripting equivalent. It provides an easy control over all key analytical parameters of the technique, novel methodological additions for augmented analyses, and facilitated visualization of the resulting CAPs and associated metrics. Although we have focused on the usefulness of CAP analysis in the resting-state setting, we also remark that nothing precludes the use of the technique in task-based investigations.

As most CAP studies to date have revealed the potential of the approach in clinical settings (Amico et al., 2014; Di Perri et al., 2018; Kaiser et al., 2019; Tuleasca et al., 2019), we sought to demonstrate the relevance of the technique in another context; i.e., rather than considering a classification problem in which two or more distinct subject populations are separated, we considered a regression task in which we attempted to explain attentional abilities within a more homogeneous population in a continuous vigilance task by means of CAPs dynamics.

We observed that of all the extracted CAPs showing coupling with the right dorsolateral prefrontal seed, the large majority either did not appear to be involved in attentional abilities, or showed positive salience weights indicating a positive impact of their expression. This is not so surprising, given that our analysis was focused on a region of the attention network in the first place. The common feature of beneficial CAPs appeared to be the coupling of an array of other regions previously pinpointed in continuous performance tasks, including the inferior parietal cortex, cerebellum lobule VI or occipital cortex (H€ager et al., 1998; Ogg et al., 2008; Tana et al., 2010). At the same time, these beneficial CAPs also barely involved coupling of other functionally distinct networks.

The one CAP for which the above reasoning does not hold is CAP₁: despite the involvement of a very diverse set of regions, it was also retrieved as beneficial for attentional performance. More precisely, contrarily to most of the others, the expression of this CAP appears to be essential at the start of a seed activation sequence: indeed, salience weights were large specifically for the entries from baseline and out-degree metrics. In other words, there is first a transition from baseline to this CAP, followed by the exit of that configuration to reach more spatially well-defined states. This involvement of short-lived periods of extensive cross-network interactions in mediating some aspects of human

Appendix

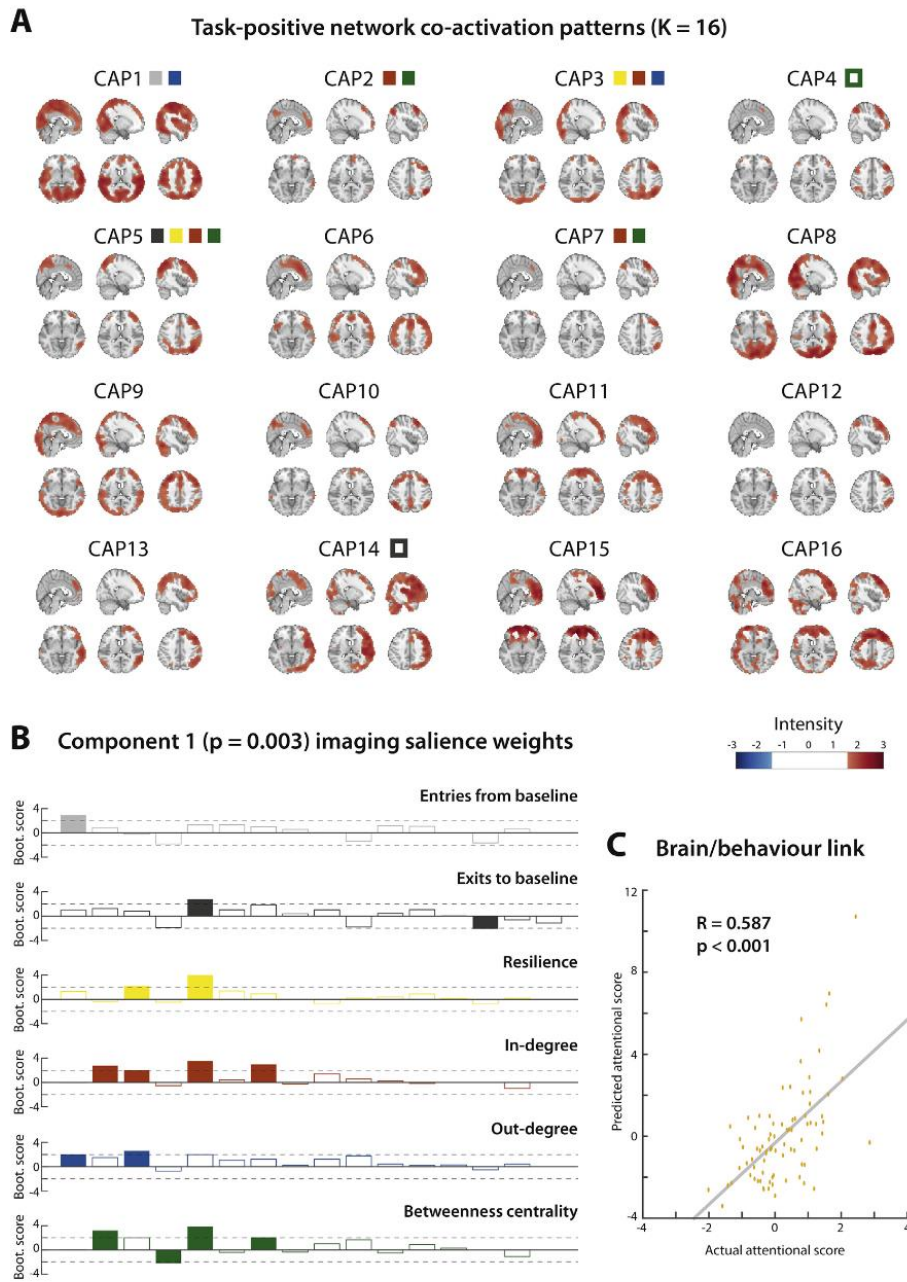


Fig. 4. Co-activation patterns, and relationship to attentional abilities. (A) The 16 obtained CAPs are plotted, with coloured rectangles symbolising the CAPs that are particularly important (bootstrap score of the associated salience weight larger than 2) for the brain/behaviour relationship as quantified from a given metric (grey: entries from baseline state, black: exits to baseline state, yellow: resilience, red: in-degree, blue: out-degree, green: betweenness centrality). Filled rectangles highlight beneficial CAPs (positive bootstrap score), while hollow rectangles depict detrimental CAPs (negative bootstrap score). (B) Salience weights across all 16 CAPs and the 6 investigated CAP metrics. The threshold bootstrap score past which a weight is considered significant is highlighted by a horizontal dashed line. Empty bars denote non-significant weights, while filled bars represent significant ones. (C) Actual attentional performance score (x-axis) versus predicted values with PLS (y-axis).

cognition has recently started to be appreciated as an insightful functional brain mechanism (Betzel et al., 2016; Fukushima et al., 2018).

The fact that all the probed metrics significantly contributed to explain attention is good evidence in favor of the temporal complexity of functional brain dynamics: instead of an instantaneous characterization or a one-frame expression of a telling functional state, what truly matters is a complex mix between how activation starts (captured by the from- baseline and to-baseline metrics),

Appendix

how transitions occur across distinct functional states (as seen from in-degree, out-degree and betweenness centrality), and how lasting a given state is (as quantified by resilience). Our characterization relates to the broad family of temporal modelling approaches, of which notable examples include the use of graph- theoretical analysis for energy landscape (Kang et al., 2019), or hidden Markov models—HMMs (Vidaurre et al., 2017; Bolton et al., 2017).

A future actual use of HMMs in CAP analysis would make it possible to not only estimate the transitory behavior across CAPs, but also the parameters governing the expression of the voxelwise patterns of BOLD signal. In addition, this full characterization of the system would enable the generation of new data, going beyond the mere computation of empirical estimates as done now. However, HMM-based approaches require extensive amounts of data to converge properly, which hinders the exploration of subject-specific properties with typical data amounts (Bolton et al., 2017). It is to bypass this issue that we instead set, in our current approach, to derive composite metrics that incorporate the information from several CAPs at once. Going back to the above example where we considered $K = 16$ CAPs, we could thus lower our amount of imaging features from $K^2 = 256$ (all of which should be estimated with an HMM) to $6 \cdot K = 96$, and this feature extraction strategy will become more and more beneficial as the number of examined CAPs increases.

Prediction of continuous performance abilities from resting-state fMRI recordings has been shown possible in previous functional connectivity work relying on second-order correlational measures across brain regions (Rosenberg et al., 2016). More recently, this characterisation has been pushed to the dynamic level by Fong et al. (2019), who showed that prediction can also be successfully achieved when temporal variability, which quantifies fluctuations in functional connectivity over the course of a scanning session, is used as a metric of interest. Our prediction accuracy is on par with the one achieved in this whole-brain analysis, despite focusing on one seed region. Interestingly, the authors described the fact that lowered temporal variability was beneficial for better attentional performance, and that many of the most important features for prediction involved an executive control brain network: this is fully consistent with our results, in which increased resilience (which can be expected to yield lowered temporal variability) of CAPs featuring executive control areas is beneficial.

In comparison to clinical applications of CAP analysis, in which between 3 and 8 CAPs are typically considered, a finer granularity was required in the present work (see Fig. 3). This is not surprising given that the regression problem at hand here is more challenging than a classification task, as we need to predict a value within a continuum. Furthermore, the functional underpinnings of inter-individual differences in cognitive abilities are likely more subtle than when comparing subjects across consciousness or disease severity levels. In fact, if a too low number of CAPs is extracted, patterns with a different cognitive relevance are averaged together as a single cluster, which impedes prediction; conversely, if a too large number of CAPs is extracted, meaningful configurations become further segmented, and statistical power is lost due to the smaller amount of frames constituting each CAP.

A limitation of our work, and of any standard CAP analysis, is that computations are performed on BOLD time courses that have not been freed from hemodynamic effects (i.e., that have not underwent deconvolution with a hemodynamic response function—HRF—estimate). The parameters of the HRF vary across brain regions (Handwerker et al., 2004), and such differences can confound functional connectivity estimates (Rangaprakash et al., 2018). However, we believe that such impacts remain

Appendix

minimal in the present analysis, since the areas located in most CAPs are also consistently found in the literature on attentional performance. Furthermore, we do not focus on the spatial patterns of the CAPs, but instead, on the transitory dynamics between them. Since the HRF also varies across subjects (Aguirre et al., 1998), we cannot rule out that our association to behaviour was partly influenced by such effects, but given the fact that our sample of subjects was distributed over a narrow age range of 10 years—a leading factor in HRF variability (D’Esposito et al., 1999), we consider this an unlikely scenario.

While the example application introduced here involved fast TR (0.72 s) data acquired in a multi-band setting, it should be remarked that the findings of CAP analyses may partly vary as a function of the employed acquisition type. Since the HRF acts as a bottleneck factor, setting a limit below which functional dynamics cannot be resolved more finely anymore even at faster acquisition paces, the influence of the TR per se may remain limited. However, another associated problematic is the differential influence of physiological rhythms on the functional data (Chen et al., 2019). This even extends to the motion time courses typically used in data preprocessing—including scrubbing as performed within TbCAPs, since additional physiology-driven components are observed at faster TRs (Power et al., 2019). All in all, we thus wish to emphasise the importance of freeing the data from such physiological impacts as well as possible using the available resources for this purpose (Glover et al., 2000; Griffanti et al., 2014; Pruim et al., 2015).

In future work, it will be interesting to examine clinical or cognitive research hypotheses at the broader focus level of more than one seed region. As alluded to above, this is already feasible with our current toolbox version, and may enable to better bridge the gap between seed-based and whole-brain analyses. We also foresee additional technical developments in the near future, such as the possibility to extract co-activation sequences (that is, series of successive fMRI volumes) rather than CAPs. Finally, we would like to encourage the motivated readers to help us in further improving our publicly accessible toolbox, so that it can become an even more multimodal package integrated with other widely used existing software.

CRedit authorship contribution statement

Thomas A.W. Bolton: Formal analysis, Writing - original draft, Conceptualization, Software, Writing - review & editing. Constantin Tuleasca: Data curation, Validation. Diana Wotruba: Validation. Gwladys Rey: Validation. Herberto Dhanis: Validation. Baptiste Gauthier: Validation. Farnaz Delavari: Validation. Elenor Morgen-roth: Validation, Writing - review & editing. Julian Gaviria: Validation. Eva Blondiaux: Validation. Lukasz Smigielski: Validation. Dimitri Van De Ville: Conceptualization, Supervision, Writing - review & editing.

Acknowledgments

Constantin Tuleasca gratefully acknowledges receipt of a ‘Young Researcher in Clinical Research Grant’ (Jeune Chercheur en Recherche Clinique) from the University of Lausanne (UNIL), Faculty of Biology and Medicine (FBM), and the Lausanne University Hospital (CHUV). In addition, the authors would like to thank Raphael Liegeois for his rereading of the manuscript, as well as the anonymous

Appendix

Reviewers who commented on this work, as they significantly contributed to improve the quality of the manuscript and of the toolbox by their insightful suggestions.

References

- Aguirre, G.K., Zarahn, E., D'esposito, M., 1998. The variability of human, BOLD hemodynamic responses. *Neuroimage* 8 (4), 360–369.
- Amico, E., Gomez, F., Di Perri, C., Vanhaudenhuyse, A., Lesenfans, D., Boveroux, P., Bonhomme, V., Brichant, J.F., Marinazzo, D., Laureys, S., 2014. Posterior cingulate cortex-related co-activation patterns: a resting state fMRI study in propofol-induced loss of consciousness. *PloS One* 9 (6), 1–9.
- Betzel, R.F., Fukushima, M., He, Y., Zuo, X., Sporns, O., 2016. Dynamic fluctuations coincide with periods of high and low modularity in resting-state functional brain networks. *Neuroimage* 127, 287–297.
- Bishop, C.M., 1999. Bayesian PCA. Curran Associates, pp. 382–388.
- Biswal, B., Zerrin Yetkin, F., Haughton, V.M., Hyde, J.S., 1995. Functional connectivity in the motor cortex of resting human brain using echo-planar MRI. *Magn. Reson. Med.* 34 (4), 537–541.
- Bolton, T.A.W., Tarun, A., Sterpenich, V., Schwartz, S., Van De Ville, D., 2017. Interactions between large-scale functional brain networks are captured by sparse coupled HMMs. *IEEE Trans. Med. Imag.* 37 (1), 230–240.
- Bolton, T.A.W., Kebets, V., Glerean, E., Zoeller, D., Li, J., Yeo, B.T.T., Caballero-Gaudes, C., Van De Ville, D., 2020. Agito ergo sum: Correlates of spatio-temporal motion characteristics during fMRI. *Neuroimage* 209, 116433.
- Bowring, A., Maumet, C., Nichols, T.E., 2019. EXploring the impact of analysis software on task fMRI results. *Hum. Brain Mapp.* 40 (11), 3362–3384.
- Caballero-Gaudes, C., Reynolds, R.C., 2017. Methods for cleaning the BOLD fMRI signal. *Neuroimage* 154, 128–149.
- Chang, C., Glover, G.H., 2010. Time-frequency dynamics of resting-state brain connectivity measured with fMRI. *Neuroimage* 50 (1), 81–98.
- Chen, J.E., Chang, C., Greicius, M.D., Glover, G.H., 2015. Introducing co-activation pattern metrics to quantify spontaneous brain network dynamics. *Neuroimage* 111, 476–488.
- Chen, J.E., Polimeni, J.R., Bollmann, S., Glover, G.H., 2019. On the analysis of rapidly sampled fMRI data. *Neuroimage* 188, 807–820.
- Damoiseau, J.S., Rombouts, S.A.R., Barkhof, F., Scheltens, P., Stam, C.J., Smith, S.M., Beckmann, C.F., 2006. Consistent resting-state networks across healthy subjects. *Proc. Natl. Acad. Sci.* 103 (37), 13848–13853.
- D'Esposito, M., Zarahn, E., Aguirre, G.K., Rypma, B., 1999. The effect of normal aging on the coupling of neural activity to the BOLD hemodynamic response. *Neuroimage* 10 (1), 6–14.
- Di, X., Biswal, B.B., Jan, 2015. Dynamic brain functional connectivity modulated by resting-state networks. *Brain Struct. Funct.* 220 (1), 37–46.

Appendix

- Di Perri, C., Amico, E., Heine, L., Annen, J., Martial, C., Larroque, S.K., Soddu, A., Marinazzo, D., Laureys, S., 2018. Multifaceted brain networks reconfiguration in disorders of consciousness uncovered by co-activation patterns. *Hum. Brain Mapp.* 39(1), 89–103.
- Fong, A.H.C., Yoo, K., Rosenberg, M.D., Zhang, S., Li, C.R., Scheinost, D., Constable, R.T., Chun, M.M., 2019. Dynamic functional connectivity during task performance and rest predicts individual differences in attention across studies. *Neuroimage* 188, 14–25.
- FoX, M.D., Greicius, M., 2010. Clinical applications of resting state functional connectivity. *Front. Syst. Neurosci.* 4, 19.
- FoX, M.D., Snyder, A.Z., Vincent, J.L., Corbetta, M., Van Essen, D.C., Raichle, M.E., 2005. The human brain is intrinsically organized into dynamic, anticorrelated functional networks. *Proc. Natl. Acad. Sci.* 102 (27), 9673–9678.
- Freeman, L.C., 1979. Centrality in social networks conceptual clarification in Hawaii nets conferences. *Soc. Network.* 1 (3), 215–239.
- Friston, K.J., 1994. Functional and effective connectivity in neuroimaging: a synthesis. *Hum. Brain Mapp.* 2 (1), 56–78.
- Friston, K.J., Holmes, A.P., Worsley, K.J., Poline, J., Frith, C.D., Frackowiak, R.S.J., 1994. Statistical parametric maps in functional imaging: a general linear approach. *Hum. Brain Mapp.* 2 (4), 189–210.
- Fukushima, M., Betzel, R.F., He, Y., de Reus, M.A., van den Heuvel, M.P., Zuo, X.,
- Sporns, O., 2018. Fluctuations between high- and low-modularity topology in time- resolved functional connectivity. *Neuroimage* 180, 406–416.
- Glover, G.H., Li, T., Ress, D., 2000. Image-based method for retrospective correction of physiological motion effects in fMRI: RETROICOR. *Magn. Reson. Med.* 44 (1), 162–167.
- Gorgolewski, K.J., Auer, T., Calhoun, V.D., Craddock, R.C., Das, S., Duff, E.P., Flandin, G.,
- Ghosh, S.S., Glatard, T., Halchenko, Y.O., et al., 2016. The brain imaging data structure, a format for organizing and describing outputs of neuroimaging experiments. *Scientific Data* 3, 160044.
- Greicius, M., 2008. Resting-state functional connectivity in neuropsychiatric disorders. *Curr. Opin. Neurol.* 21 (4), 424–430.
- Griffanti, L., Salimi-Khorshidi, G., Beckmann, C.F., Auerbach, E.J., Douaud, G., Sexton, C.E., Zsoldos, E., Ebmeier, K.P., Filippini, N., Mackay, C.E., et al., 2014. ICA- based artefact removal and accelerated fMRI acquisition for improved resting state network imaging. *Neuroimage* 95, 232–247
- Gur, R.C., Richard, J., Huggett, P., Calkins, M.E., Macy, L., Bilker, W.B., Brensinger, C., Gur, R.E., 2010. A cognitive neuroscience-based computerized battery for efficient measurement of individual differences: standardization and initial construct validation. *J. Neurosci. Methods* 187 (2), 254–262
- Haeger, F., Volz, H., Gaser, C., Mentzel, H., Kaiser, W.A., Sauer, H., 1998. Challenging the anterior attentional system with a continuous performance task: a functional magnetic resonance imaging approach. *Eur. Arch. Psychiatr. Clin. Neurosci.* 248 (4), 161–170.
- Handwerker, D.A., Ollinger, J.M., D’Esposito, M., 2004. Variation of BOLD hemodynamic responses across subjects and brain regions and their effects on statistical analyses. *Neuroimage* 21 (4), 1639–1651.
- Holmes, A.J., Hollinshead, M.O., O’Keefe, T.M., Petrov, V.I., Fariello, G.R., Wald, L.L.,

Appendix

- Fischl, B., Rosen, B.R., Mair, R.W., Roffman, J.L., et al., 2015. Brain Genomics Superstruct Project initial data release with structural, functional, and behavioral measures. *Scientific Data* 2, 150031.
- Hutchison, R.M., Womelsdorf, T., Allen, E.A., Bandettini, P.A., Calhoun, V.D., Corbetta, M., Della Penna, S., Duyn, J.H., Glover, G.H., Gonzalez-Castillo, J., Handwerker, D.A., Keilholz, S., Kiviniemi, V., Leopold, D.A., de Pasquale, F., Sporns, O., Walter, M., Chang, C., 2013. Dynamic functional connectivity: promise, issues, and interpretations. *Neuroimage* 80, 360–378.
- Kaiser, R.H., Kang, M.S., Lew, Y., Van Der Feen, J., Aguirre, B., Clegg, R., Goer, F., Esposito, E., Auerbach, R.P., Hutchison, R.M., et al., 2019. Abnormal frontoinsula- default network dynamics in adolescent depression and rumination: a preliminary resting-state co-activation pattern analysis. *Neuropsychopharmacology* 44, 1604–1612.
- Kang, J., Pae, C., Park, H., 2019. Graph-theoretical analysis for energy landscape reveals the organization of state transitions in the resting-state human cerebral cortex. *PloS One* 14 (9), 0222161.
- Kriegeskorte, N., Simmons, W.K., Bellgowan, P.S.F., Baker, C.I., 2009. Circular analysis in systems neuroscience: the dangers of double dipping. *Nat. Neurosci.* 12 (5), 535–540.
- Krishnan, A., Williams, L.J., McIntosh, A.R., Abdi, H., 2011. Partial least squares (PLS) methods for neuroimaging: a tutorial and review. *Neuroimage* 56 (2), 455–475.
- Liu, X., Chang, C., Duyn, J.H., 2013. Decomposition of spontaneous brain activity into distinct fMRI co-activation patterns. *Front. Syst. Neurosci.* 7, 1–11.
- Liu, X., Duyn, J.H., 2013. Time-varying functional network information extracted from brief instances of spontaneous brain activity. *Proc. Natl. Acad. Sci.* 110 (11), 4392–4397.
- Liu, X., Zhang, N., Chang, C., Duyn, J.H., 2018. Co-activation patterns in resting-state fMRI signals. *Neuroimage* 180, 485–494.
- Logothetis, N.K., Pauls, J., Augath, M., Trinath, T., Oeltermann, A., 2001. Neurophysiological investigation of the basis of the fMRI signal. *Nature* 412 (6843), 150–157.
- McIntosh, A.R., Lobaugh, N.J., 2004. Partial least squares analysis of neuroimaging data: applications and advances. *Neuroimage* 23, 250–263.
- Monti, S., Tamayo, P., Mesirov, J., Golub, T., 2003. Consensus clustering: a resampling-based method for class discovery and visualization of gene expression microarray data. *Mach. Learn.* 52 (1), 91–118.
- Murphy, K., Fox, M.D., 2017. Towards a consensus regarding global signal regression for resting state functional connectivity MRI. *Neuroimage* 154, 169–173.
- Nooner, K.B., Colcombe, S., Tobe, R., Mennes, M., Benedict, M., Moreno, A., Panek, L., Brown, S., Zavitz, S., Li, Q., et al., 2012. The NKI-Rockland sample: a model for accelerating the pace of discovery science in psychiatry. *Front. Neurosci.* 6, 152.
- Ogg, R.J., Zou, P., Allen, D.N., Hutchins, S.B., Dutkiewicz, R.M., Mulhern, R.K., 2008. Neural correlates of a clinical continuous performance test. *Magn. Reson. Imag.* 26 (4), 504–512.
- Power, J.D., Barnes, K.A., Snyder, A.Z., Schlaggar, B.L., Petersen, S.E., 2012. Spurious but systematic correlations in functional connectivity MRI networks arise from subject motion. *Neuroimage* 59 (3), 2142–2154.
- Power, J.D., Cohen, A.L., Nelson, S.M., Wig, G.S., Barnes, K.A., Church, J.A., Vogel, A.C., Laumann, T.O., Miezin, F.M., Schlaggar, B.L., et al., 2011. Functional network organization of the human brain. *Neuron* 72 (4), 665–678.

Appendix

- Power, J.D., Lynch, C.J., Silver, B.M., Dubin, M.J., Martin, A., Jones, R.M., 2019. Distinctions among real and apparent respiratory motions in human fMRI data. *Neuroimage* 201, 116041.
- Preti, M.G., Bolton, T.A.W., Van De Ville, D., 2017. The dynamic functional connectome: state-of-the-art and perspectives. *Neuroimage* 160, 41–54.
- Pruim, R.H., Mennes, M., van Rooij, D., Llera, A., Buitelaar, J.K., Beckmann, C.F., 2015. ICA-AROMA: a robust ICA-based strategy for removing motion artifacts from fMRI data. *Neuroimage* 112, 267–277.
- Rangaprakash, D., Wu, G., Marinazzo, D., Hu, X., Deshpande, G., 2018. Hemodynamic response function (HRF) variability confounds resting-state fMRI functional connectivity. *Magn. Reson. Med.* 80 (4), 1697–1713.
- Rosenberg, M.D., Finn, E.S., Scheinost, D., Papademetris, X., Shen, X., Constable, R.T., Chun, M.M., 2016. A neuromarker of sustained attention from whole-brain functional connectivity. *Nat. Neurosci.* 19 (1), 165.
- Rubinov, M., Sporns, O., 2010. Complex network measures of brain connectivity: uses and interpretations. *Neuroimage* 52 (3), 1059–1069.
- Satterthwaite, T.D., Wolf, D.H., Loughhead, J., Ruparel, K., Elliott, M.A., Hakonarson, H., Gur, R.C., Gur, R.E., 2012. Impact of in-scanner head motion on multiple measures of functional connectivity : relevance for studies of neurodevelopment in youth. *Neuroimage* 60 (1), 623–632.
- Senbabaoglu, Y., Michailidis, G., Li, J.Z., 2014. Critical limitations of consensus clustering in class discovery. *Sci. Rep.* 4, 6207.
- Shirer, W.R., Ryali, S., Rykhlevskaia, E., Menon, V., Greicius, M.D., 2012. Decoding subject-driven cognitive states with whole-brain connectivity patterns. *Cerebr. Cortex* 22 (1), 158–165.
- Smith, S.M., Beckmann, C.F., Andersson, J., Auerbach, E.J., Bijsterbosch, J., Douaud, G., Duff, E., Feinberg, D.A., Griffanti, L., Harms, M.P., et al., 2013. Resting-state fMRI in the human connectome project. *Neuroimage* 80, 144–168.
- Tagliazucchi, E., Balenzuela, P., Fraiman, D., Chialvo, D.R., 2012. Criticality in large-scale brain fMRI dynamics unveiled by a novel point process analysis. *Front. Physiol.* 3, 1–12.
- Tana, M.G., Montin, E., Cerutti, S., Bianchi, A.M., 2010. EXploring cortical attentional system by using fMRI during a continuous performance test. *Comput. Intell. Neurosci.* 2010, 329213.
- Tuleasca, C., Bolton, T.A.W., R'egis, J., Najdenovska, E., Witjas, T., Girard, N., Delaire, F., Vincent, M., Faouzi, M., Thiran, J., et al., 2019. Normalization of aberrant pretherapeutic dynamic functional connectivity of extrastriate visual system in patients who underwent thalamotomy with stereotactic radiosurgery for essential tremor: a resting-state functional MRI study. *J. Neurosurg.* 1, 1–10.
- van den Heuvel, M.P., Hulshoff Pol, H.E., 2010. EXploring the brain network: a review on resting-state fMRI functional connectivity. *Eur. Neuropsychopharmacol* 20 (8), 519–534.
- Van Dijk, K.R., Sabuncu, M.R., Buckner, R.L., 2012. The influence of head motion on intrinsic functional connectivity MRI. *Neuroimage* 59 (1), 431–438.
- Van Essen, D.C., Smith, S.M., Barch, D.M., Behrens, T.E.J., Yacoub, E., Ugurbil, K., 2013. The WU-Minn human connectome project: an overview. *Neuroimage* 80, 62–79.
- Vidaurre, D., Smith, S.M., Woolrich, M.W., 2017. Brain network dynamics are hierarchically organized in time. *Proc. Natl. Acad. Sci.* 114 (48), 201705120.

Appendix

Zhuang, X., Walsh, R.R., Sreenivasan, K., Yang, Z., Mishra, V., Cordes, D., 2018.

Incorporating spatial constraint in co-activation pattern analysis to explore the dynamics of resting-state networks: an application to Parkinson's disease. *Neuroimage* 172, 64–84.

Appendix

6.2 Thought Consciousness and source monitoring depend on robotically controlled sensorimotor conflicts and illusory states

Andrea Serino ^{1,2,3, §}, Polona Pozeg ^{1,2, §}, Fosco Bernasconi ^{1,2}, Marco Solcà ^{1,2}, Masayuki Hara ⁴, Pierre Progin ^{5,6}, Giedre Stripeikyte ^{1,2}, **Herberto Dhanis** ^{1,2}, Roy Salomon ^{1,2,7}, Hannes Bleuler ⁸, Giulio Rognini ^{1,2,8, §§}, Olaf Blanke ^{1,2,9, §§,*}

Affiliations

¹ Laboratory of Cognitive Neuroscience, Brain Mind Institute, Faculty of Life Sciences, Swiss Federal Institute of Technology (EPFL), Geneva, Switzerland

² Center for Neuroprosthetics, School of Life Sciences, Campus Biotech, Swiss Federal Institute of Technology, Ecole Polytechnique Fédérale de Lausanne (EPFL), 1012 Geneva, Switzerland

³ MySpace Lab, Department of Clinical Neurosciences, University Hospital of Lausanne (CHUV), Lausanne, Switzerland

⁴ Control Engineering Laboratory, Graduate School of Science and Engineering, Saitama University, Saitama, 338-8570, Japan

⁵ Center for Psychiatric Neuroscience, Centre Hospitalier Universitaire Vaudois (CHUV), University of Lausanne (UNIL), Lausanne, Switzerland

⁶ Service of General Psychiatry, Centre Hospitalier Universitaire Vaudois (CHUV), University of Lausanne (UNIL), Lausanne, Switzerland

⁷ Gonda Brain Research Center, Bar-Ilan University, Ramat Gan, Israel

⁸ Laboratory of Robotic Systems, Swiss Federal Institute of Technology (EPFL), Lausanne, Switzerland

⁹ Service de Neurologie, University Hospital Geneva, Geneva, Switzerland

[§] Both authors contributed equally

^{§§} Both authors contributed equally

Corresponding Author

Prof. Olaf Blanke

Bertarelli Chair in Cognitive Neuroprosthetics
Center for Neuroprosthetics & Brain Mind Institute
School of Life Sciences
Ecole Polytechnique Fédérale de Lausanne (EPFL)
Campus Biotech, H4.3
Ch. des Mines 9
CH-1202 Geneva
E-mail: olaf.blanke@epfl.ch
Tel: +41 (0)21 693 69 21

Appendix

Abstract

Thought insertion (TI) is characterized by the experience that certain thoughts, occurring in one's mind, are not one's own, but the thoughts of somebody else and suggestive of a psychotic disorder. We report a robotics-based method able to investigate the behavioral and subjective mechanisms of TI in healthy participants. We used a robotic device to alter body perception by providing online sensorimotor stimulation, while participants performed cognitive tasks implying source monitoring of mental states attributed to either oneself or another person. Across several experiments, conflicting sensorimotor stimulation reduced the distinction between self- and other-generated thoughts and was, moreover, associated with the experimentally generated feeling of being in the presence of an alien agent and subjective aspects of TI. Introducing a new robotics-based approach that enables the experimental study of the brain mechanisms of TI, these results link TI to predictable self-other shifts in source monitoring and specific sensorimotor processes.

Appendix

Introduction

Thought insertion (TI) is one of the most enigmatic psychiatric symptoms and is characterized by the experience that certain thoughts, occurring in one's mind, are not one's own, but rather the thoughts of somebody else. TI violates basic intuitions about consciousness (i.e., Who else than me could possibly have access to my thoughts?) and has fascinated clinicians, scientists, philosophers, and laymen alike. TI is often reported by patients with schizophrenia and other psychotic disorders and may rarely occur in healthy individuals (Johns et al., 2004). TI is classified as the so-called first-rank symptom, implying that a regular occurrence is suggestive of a psychotic disorder (Schneider, 1959).

A long-standing question in psychiatric and cognitive neuroscience has been how the brain generates TI and on which brain mechanisms it depends. One prominent postulation is that first-rank symptoms, including TI, arise from a deficit comparable to those of conscious control for overt actions, that is, a deficit of source monitoring (Feinberg, 1978; Frith, 1987; Ford and Mathalon, 2004) and related sensorimotor mechanisms. This proposal is substantiated by converging behavioral, brain imaging, and electrophysiological evidence in patients with schizophrenia (Ford and Mathalon, 2004; Shergill et al., 2005, 2014) and healthy subjects (Weiskrantz et al., 1971; Shergill et al., 2003; Bays et al., 2005, 2006), but has so far targeted only conscious control of overt actions or auditory verbal hallucinations (i.e., alien voices, Hoffman, 1986). Accordingly, the importance of source monitoring, self-related processes, and the link of TI to conscious monitoring of overt actions, remains poorly understood. Although some authors have investigated the mechanisms related to TI using different cognitive manipulations (Walsh et al., 2015; Sugimori et al., 2011) (see also Stephens and Graham, 2000; Martin and Pacherie, 2013; Gallagher, 2004a; 2004b; Vicente, 2014), research on TI and related cognitive processes has been hampered by the lack of empirical techniques in healthy subjects to probe TI and investigate associated behavioral changes in a more controlled fashion. Accordingly, the mechanisms of TI, and how they potentially depend on conscious control for overt actions and covert mental activity, remain unknown. To provide empirical evidence about the interaction between the sensorimotor control of actions and covert mental activity in potentially generating TI, here we applied a robotic device that allowed us to interfere in a specific and controlled way with sensorimotor processing (known to alter source monitoring), while participants performed repetitive cognitive tasks.

Our recently developed robotic system consists of two robots and has previously allowed us to experimentally alter own body perception and, importantly, is able to induce illusory mental states mimicking psychosis-related symptoms, in a controlled manner in healthy subjects (Blanke et al., 2014). During the procedure participants are asked to perform repeated poking movements, through a front robot (i.e., placed in front of participants) (Figure 1) and replicated by a back robot (i.e., placed behind the participants), resulting in controlled tactile stimulation on the participants' back based on their own movements (synchronous stimulation). Blanke et al. (2014) demonstrated that if a temporal delay is introduced between the participants' movements and the tactile stimulation delivered on their back (i.e., asynchronous sensorimotor stimulation) healthy participants experience an illusory alteration of their mental state characterized by passivity and loss of agency, as well as being in the presence of somebody else (feeling of an alien presence) (feeling of a presence [FoP]).

Appendix

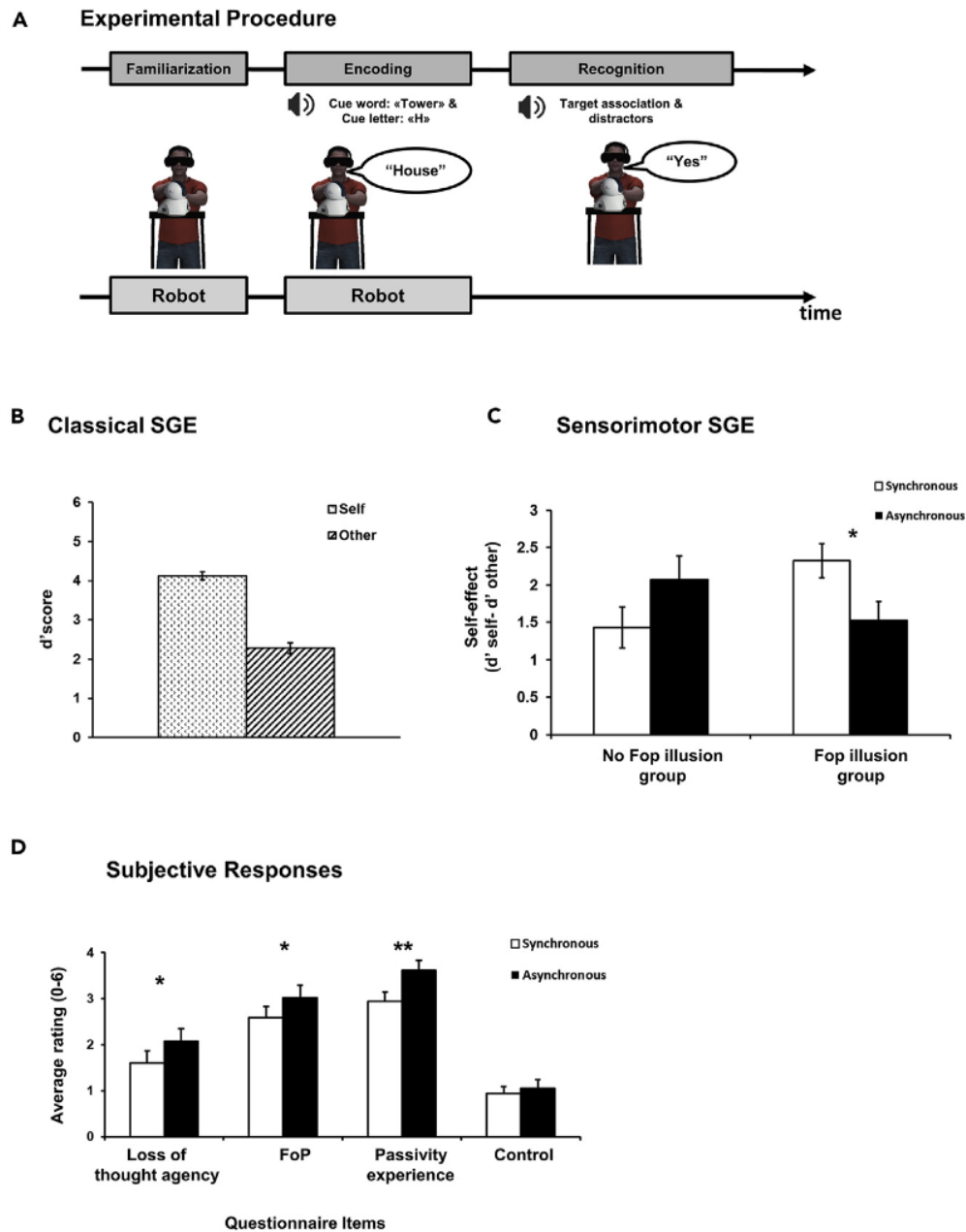


Figure 1. Thought generation (experiment 1)

(A) Experimental procedure for Experiment 1. During encoding, participants operate the robotic system in synchronous or asynchronous mode, followed by the memory recognition phase. Participants answered whether they had generated (active condition) or heard (passive condition) the word.

(B) Classical SGE (d') was higher in the active versus passive conditions.

(C) Only individuals experiencing the FoP had significantly less self-advantage (sensorimotor SGE; $d'_{active} - d'_{passive}$) in the asynchronous when compared with the synchronous condition (error bars standard error of mean).

(D) Participants reported stronger FoP, passivity experiences, and loss of thought agency during the asynchronous versus synchronous condition. Error bars show standard errors of the mean. * $p < 0.05$, ** $p < 0.01$.

In four separate experiments, we investigated whether source monitoring for internal thoughts depends on (1) sensorimotor stimulation and on (2) the level of FoP while exposing our participants to asynchronous and synchronous (i.e., control condition) robotic stimulation. Importantly, for the present experiments, previous work has shown that participants are able to carry out different covert cognitive paradigms while they are also actuating the robotic system and hence receive sensorimotor

Appendix

stimulation (i.e., Salomon et al., 2020; Faivre et al., 2020; Orepic et al., 2020). In the present experiments, in Experiment 1, we tested the effects of robotic stimulation and FoP on source monitoring in a memory task by exploiting the so-called self-generation effect (SGE) and in Experiment 2 in a new task developed to assess thought numerosity (during a verbal fluency task). In Experiment 3, we investigated whether sensorimotor stimulation and the thought numerosity paradigm were associated with explicit changes in subjective thought experience. In a final control experiment, Experiment 4, we excluded that the observed effects were due to a generic reduction of attentional resources during asynchronous stimulation (by using a classical working memory task). Across these four experiments we demonstrate systematic behavioral and subjective changes in source monitoring suggestive of TI while participants performed different mental operations, which depend on online conflicting sensorimotor stimulation and the level of experienced FoP. We discuss the importance of robotics and sensorimotor processes for the understanding of cognitive thought processes, including thought agency as well as abnormal and clinically relevant TI.

Results

Robotically-induced sensorimotor conflict induce FoP and alter source monitoring

In Experiment 1, we used a robotic system (Blanke et al., 2014; Salomon et al., 2020; Faivre et al., 2020; Orepic et al., 2020; Hara et al., 2011) (Figure S1) and exposed a group of healthy participants to repetitive sensorimotor stimulation that induces the FoP in a controlled way (see below) while they simultaneously performed a mental source monitoring task (Experiment 1). In this paradigm, inducing the so-called SGE (Slamecka and Graf, 1978; Transparent methods), the participants were either presented with a list of words (passive condition) or they were asked to generate their own words (active condition) within a given set of rules. During an encoding phase, participants were asked to memorize both the self-generated and the passively heard words. When tested in a subsequent recognition phase, participants typically remember more self-generated than externally presented (heard) items, i.e., SGE. To avoid ceiling and floor effects in the recognition task for self-generated words, only the data of participants who generated more than 50% of expected associations (at least 18 words) and who performed above chance in the recognition task were included to the analysis.

Importantly for the present investigation, our participants additionally performed repetitive tapping movements with both hands to operate the front robot, which was combined with a second robot providing tactile feedback to their back (see Transparent methods, for more detail). In two conditions, tactile feed-back was delivered either synchronously with their movement (synchronous control condition) or with a delay (of 500 ms; asynchronous condition) that, critically, we previously showed induces the FoP in healthy participants. In the first part of Experiment 1, while participants were using our robotic system, we asked them to carry out the standard procedure to measure the SGE. Previous work on the sense of agency for overt actions (and its link to self or source monitoring processes) has typically exposed subjects to different sensorimotor conditions, by varying the spatiotemporal contingencies between actions and associated sensory feedback or by measuring consequences in terms of sensory attenuation or motor adaptations (Shergill et al., 2003; Bays et al., 2005, 2006; Blakemore et al., 1998, 2000; Wolpert and Ghahramani, 2020). As indicated above, the SGE is a well-known memory effect, characterized by better recognition for words that are self-

Appendix

generated (active condition) versus words that are only heard and generated by another person (passive condition, Faivre et al., 2020) (Figure 1A; Transparent methods). Here, to study the relation between thought-related source monitoring and sensorimotor processing, we tested whether the magnitude of the SGE (i.e., the difference between recognition for self-generated versus generated-by-another words) was affected by the synchronous-asynchronous manipulation and the associated robotically induced FoP. Participants used the robotic system, either in the synchronous or asynchronous condition, during the word encoding session, i.e., while they were either generating or listening to words. The SGE for word recognition was tested immediately afterward. We hypothesized that if asynchronous stimulation induces the FoP, it might also decrease source monitoring for self-generated concurrent mental operations, by decreasing a classical self-effect such as the SGE.

As expected, we found a classical SGE (calculated as a recognition difference in d' between active and passive conditions), with significantly better recognition for actively (self) versus passively (other) generated words (Figure 1B) (self: $M = 4.12$, $SD = 0.45$; other: $M = 2.28$, $SD = 0.65$; $F(1,20) = 180.86$, $p < 0.0001$), confirming that participants better remembered words for which they have been the agents, when compared with words they passively heard. Critically, the SGE was modulated by the sensorimotor conditions (asynchronous versus synchronous), and this depended on the FoP intensity (calculated as the difference between FoP ratings in the synchronous and asynchronous condition) (significant interaction between stimulation condition and FoP intensity scores used as a covariate; $F(1,20) = 6.95$; $p = 0.016$) (Figure 1C). To better illustrate how sensorimotor stimulation inducing the FoP effect differently affected recognition in the active and passive conditions, we divided the sample in two groups accordingly to their FoP ratings and directly compared the SGE between participants who did and who did not experience to be in the presence of an alien agent. There was a significant interaction between sensorimotor condition and FoP group ($F(1,20) = 7.217$, $p = 0.014$): the SGE (i.e., difference between active and passive conditions) was lower in the FoP inducing asynchronous (versus synchronous) condition, but only in participants experiencing the FoP (FoP group, synchronous: $M = 2.33$, $SD = 0.75$; asynchronous: $M = 1.53$, $SD = 0.83$) (Figure 1C). This was not the case in the other group of participants (No-FoP group, synchronous: $M = 1.43$; $SD = 0.90$; asynchronous: $M = 2.07$, $SD = 1.04$). In other words, when the robotically applied sensorimotor conflict induced the experience to be in the presence of an alien agent (FoP), the SGE, an overt behavioral advantage in the ability to remember self-generated (active condition) versus other-generated words (passive condition), was reduced. Importantly this effect was not due to a general interference on memory performance due to the robotic stimulation or to the induced FoP, as there was no main effect of sensorimotor stimulation ($p = 0.58$) or a Stimulation \times FoP interaction ($p = 0.28$) on the performance in word recognition in general. This is an important control, excluding that the differences found in the SGE depend on generic differences in distraction or divided attention between the two sensorimotor conditions. Finally, we note that the present results cannot be due to differences in motor patterns spontaneously adopted by participants during the synchronous versus the asynchronous stimulation condition. Indeed, data collected with the same robotic system show that there is no difference in the quantity of poking movements performed in the two conditions, and that there is no link between movement characteristics and the induced FoP (Bernasconi et al., 2020).

Appendix

To summarize, data from Experiment 1 show that the present sensorimotor conflicts induce selective behavioral changes in the SGE that tap into the brain's source monitoring processes (Figure 1C). Importantly, this SGE decrease in our participants' capacity to better remember self-generated versus other-generated words depends on the degree of feeling of an alien presence as induced by robotic stimulation (Figure 1D) and only in the conflicting asynchronous condition.

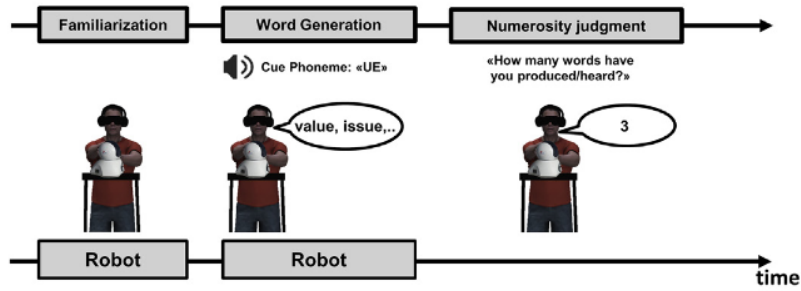
Thought numerosity is associated with source monitoring and the feeling of an alien presence

Blanke et al. (2014) demonstrated that the FoP, induced by the robotic stimulation in the asynchronous condition, was also associated with a change in how many people participants perceived to be close to them during sensorimotor stimulation, such that participants perceived additional people to be present during the FoP-inducing asynchronous condition. Here we asked whether a similar change in numerosity judgments also occurs for the number of concurrent internal thoughts participants hold in their mind. This was also motivated because TI is not only characterized by the experience that certain thoughts, occurring in one's mind, are not one's own thoughts (loss of thought agency), but also by the sensation (or positive symptom) that the thoughts in one's mind are the thoughts of a different, alien and additional, person (i.e., TI proper, Stephens and Graham, 2000; Martin and Pacherie, 2013). A lack of self-other discrimination or decrease in source monitoring as found in Experiment 1 is therefore not sufficient to account for TI that is also characterized by TI proper, because the former does not include a positive mental element characterized by the conscious attribution of one's thoughts to another additional agent. Moreover, the lack of thought agency without TI proper may also occur in healthy subjects, as is the case during unbidden thoughts (Stephens and Graham, 2000; Martin and Pacherie, 2013; Koehler, 1979), whereas TI proper has, to the best of our knowledge, not been reported in healthy subjects.

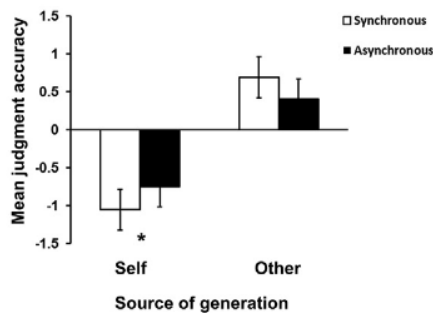
In Experiment 2 we investigated whether we can obtain a behavioral index for alienated thoughts similar to TI proper, which is an index for additional-inserted number of thoughts in healthy participants, and how this depends on the FoP. Blindfolded participants operated the same robotic system, while simultaneously performing a verbal (phonetic) fluency task (Slamecka and Graf, 1978). With the aim to observe changes in overt behavior that are associated with TI proper, we adapted a verbal fluency task and asked a group of participants to estimate the number of words that they have either generated themselves (active condition) or listened to (passive condition), while operating the robotic sensorimotor system in either the synchronous or asynchronous condition. In the active condition, a starting phoneme was played to participants through headphones and they were instructed to generate as many words starting with the specified phoneme as they could in a given time period (phonetic fluency task), which randomly varied between 15 and 30 s. Immediately afterward, each participant estimated how many words he or she had generated. In the passive conditions, the participant listened to a list of words (of 6–10 words, randomized) (Figure 2A; Transparent methods). To prevent participants from simply counting the words in the passive condition, and to avoid strong differences in cognitive load required between

Appendix

A Experimental Procedure



B Thought numerosity



C Thought numerosity judgments correlate with FoP

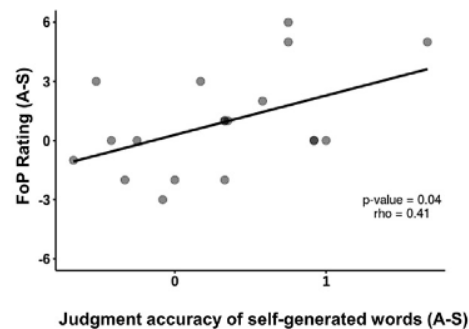


Figure 2. Thought numerosity task (experiment 2)

(A) While operating the robotic system in synchronous or asynchronous mode participants performed the thought numerosity task (either active, self-generating, or passive conditions) ([Transparent methods](#)).

(B) Thought numerosity judgments are shown. Participants showed a general suppression of numerosity judgments for self-generated words (active conditions). Crucially, this self-suppression was reduced during asynchronous versus synchronous condition. There was no such change for other-generated words (passive condition).

(C) Correlation analysis shows a significant positive correlation between the magnitude of numerosity judgment suppression and the differential FoP score. Error bars show standard errors of the mean. * $p < 0.05$.

the two conditions, they were asked to determine whether each word they heard contained a given phoneme, specified at the beginning of each trial. To obtain a measure of how well subjects are able to estimate the number of “thoughts in their mind” (i.e., thought numerosity), we subtracted the actual number of produced (active condition) or passively heard words (passive condition) from the estimated number of words. We predicted that sensorimotor stimulation should (1) differently impact word numerosity, but specifically in the active self-generating condition (i.e., more thoughts as quantified through word numerosity judgments) and that (2) this should again (as in Experiment 1) be related to the strength of the robotically induced FoP.

We found that participants underestimated the number of self-generated words ($M = -0.90$, $SD = 1.13$) when compared with words generated by another agent ($M = 0.55$, $SD = 1.11$; main effect active-passive: $F(1,18) = 23.306$, $p < 0.0001$). Critically, this self-suppression effect depended on sensorimotor stimulation (active-passive by sensorimotor condition interaction: $F(1,18) = 7.274$, $p = 0.015$), as the number of estimated words in the active conditions differed in the asynchronous ($M = -0.75$, $SD = 1.16$) versus synchronous condition ($M = -1.05$, $SD = 1.17$; $t(18) = 2.192$, $p = 0.042$). This was not observed when words were processed in the passive conditions (synchronous: $M = 0.69$, $SD = 1.20$; asynchronous: $M = 0.41$, $SD = 1.14$; $t(18) = 1.668$, $p = 0.113$) (Figure 2B), showing that these

Appendix

behavioral changes are not related to differences in attentional resources between the sensorimotor conditions or between the passive versus active condition.

We next tested whether this effect, that jointly depends on sensorimotor stimulation (asynchronous-synchronous difference) and source monitoring (active-passive difference), is also associated with the FoP. This was confirmed by the finding that the asynchronous-synchronous difference for the numerosity judgment of actively generated words correlated positively with the FoP intensity ($\rho = 0.41$, $p = 0.04$) (Figure 2C). That is, the more intense a participant experienced the FoP, the more her self-suppression effect in thought numerosity judgments was reduced in the asynchronous (when compared with the synchronous) condition, that is perceived numerosity of self-generated words became more similar to other-generated words.

Additional analyses excluded that these effects were due to generic differences in attentional resources or cognitive load between experimental conditions. There was neither a difference in the total number of generated words in the active condition ($M = 7.95$, $SD = 2.02$) and the number of words where the correct phoneme was identified in the passive condition ($M = 8.11$, $SD = 0.33$; $F(1,18) = 0.115$, $p = 0.738$), or between both sensorimotor conditions (synchronous: $M = 8.18$, $SD = 1.18$; asynchronous: $M = 7.88$, $SD = 0.89$; $F(1,18) = 3.079$, $p = 0.096$), nor was there an interaction between the source (active-passive) and sensorimotor stimulation ($F(1,18) = 0.944$, $p = 0.344$). These effects were also not modulated by the experienced FoP, as when adding FoP ratings as a covariate, no main effects or interactions emerged (all p values > 0.35 ; see also Transparent methods). This is an important control and, extending the results obtained for Experiment 1, excludes that the differences in the estimated number of words depended on general differences in distraction, divided attention, or task difficulty between the two sensorimotor conditions.

To summarize, these data reveal a robotically induced reduction of thought-related source monitoring characterized by a reduced ability to discriminate mental processes representing self-generated thoughts from those generated by others, making thought numerosity judgments more similar for words that were either actively generated or passively heard, independently of differences in cognitive load between the present experimental conditions. Importantly, the direction of the self-suppression effect suggests that perceived thought numerosity in the asynchronous active condition (when compared with the synchronous active condition) is shifted toward performance in the passive conditions, i.e., in conditions during which participants judge items generated by another person. This was further corroborated by linking this shift in performance to the experimental induction of being in the presence of an alien agent (FoP), because self-generated words were perceived as more similar to other-generated words in the FoP-inducing asynchronous condition and because the self-suppression effect correlated positively with FoP intensity. Accordingly, the number of self-generated words were perceived as higher and more similar to the number of other-generated words, selectively in the FoP-inducing asynchronous condition, suggesting that under these conditions additional and alien-like thoughts were inserted into the minds of our participants (TI proper), compatible with previous findings on the perceived number of alien people (Blanke et al., 2014).

Appendix

Sensorimotor mental state related to TI depends on the feeling of an alien presence and sensorimotor stimulation

We finally sought to provide additional evidence whether the experimental conditions leading to the changes in overt behavior in Experiment 2 are associated with changes in subjective TI and whether this depends on processes of source monitoring and the FoP. To this aim in Experiment 3, we asked a new group of participants to perform the verbal fluency task (active condition as in Experiment 2), while operating the robotic system in either the synchronous or asynchronous condition (see Transparent methods). At the beginning of each condition, they heard a French phoneme through headphones, and were then asked to generate as many words as they could, starting with the specified phoneme within 3 min (phonetic fluency task, Lezak et al., 1995). At the end of each condition, they were asked to rate the items on a questionnaire referring to their thought process during the task (Figure 3A). The questionnaire was based on previous TI literature (Miller et al., 1999; Schultze-Lutter et al., 2007) and contained a total of twelve items, with six items assessing TI and other aspects of thought consciousness, as well as six control items (Table S1).

Both sensorimotor conditions were then repeated in randomized order (without the verbal fluency task) followed by the FoP questionnaire as used in the previous experiments (Transparent methods). We pre- dicted that experimental TI and related aspects of thought consciousness would be stronger during asyn- chronous versus synchronous sensorimotor stimulation and that it would be associated with the experience of an alien presence (FoP).

Accordingly, results showed that that sensorimotor stimulation affected thought-related items, but not control items, and that this effect depended on the FoP strength as induced by the asynchronous stimulation. Indeed, there was a significant interaction between the type of question (thoughts experience, control), sensorimotor stimulation (synchronous, asynchronous), and FoP score ($F(1,17) = 7.49$, $p = 0.011$, $h^2 = 0.30$). Further analysis, run on thought experience questions only, showed a marginally significant stimulation 3 FoP interaction ($F(1,14) = 4.32$, $p = 0.05$; $h^2 = 0.19$), suggesting that the sensorimotor stimulation conditions differently affected subjects responses, as a function of whether they did or did not perceive the FoP. When analyzing individual questions, the sensorimotor X Question (Q1, Q3, Q7, Q8, Q10, and Q11) 3 FoP interaction was significant ($F(5,85) = 4.60$, $p < 0.001$; $h^2 = 0.19$), indicating that the effect of sensorimotor stimulation was stronger for some key experimental questions assessing different aspects of thoughts experience. Question-by-question analysis then revealed that, while performing the verbal fluency task, our participants reported mild experiences of thought insertion (“It seemed as if the robot put certain thoughts in my mind”) and that their thoughts were manipulated (“It seemed as if the robot influenced some of my thoughts”). Importantly, as predicted, experimental TI and influence were stronger in the asynchronous than in the synchronous condition (thought influence; asynchronous: $M = 3.33$, $SD = 1.64$, synchronous: $M = 1.89$, $SD = 1.49$; Wilcoxon signed-rank test: $Z = 2.34$, $p = 0.01$) (TI; asynchronous: $M = 2.00$, $SD = 1.41$, synchronous: $M = 1.61$, $SD = 1.38$; Wilcoxon signed-rank test: $Z = 2.11$, $p = 0.03$; asynchronous: $M = 2.5$, $SD = 1.71$, synchronous: $M = 1.67$, $SD = 1.15$; Wilcoxon signed-rank test: $Z = -1.91$, $p = 0.03$) (Figure 3B; Table S1). As expected, participants also gave higher ratings for the FoP in the asyn- chronous ($M = 3.95$; $SD = 2.07$) versus synchronous condition ($M = 2.56$; $SD = 2.06$) (Wilcoxon signed-rank test: $Z = -2.69$, $p = 0.005$) and for passivity experiences (asynchronous: $M = 4.5$, $SD = 1.61$; synchronous: $M = 2.77$, $SD = 1.69$; Wilcoxon signed-rank test: $Z = -2.57$, $p = 0.007$; Transparent methods). Further analysis revealed that the strength of thought insertion and thought influencing

Appendix

positively correlated with the intensity of the FoP (thought insertion: $\rho = 0.56$, $p = 0.01$; thought influencing: $\rho = 0.69$, $p = 0.001$) (Figure 3C). These selective effects were absent for control questions. We only observed a significant effect of question ($F(5,80) = 5.41$, $p < 0.001$, $h^2 = 0.25$), showing that participants gave different ratings to the different items; however, these ratings did not differ as a function of sensorimotor stimulation and were not influenced by the FoP effect, as no other main effect or interaction was significant (all p values $> .13$). These results rule out a possible

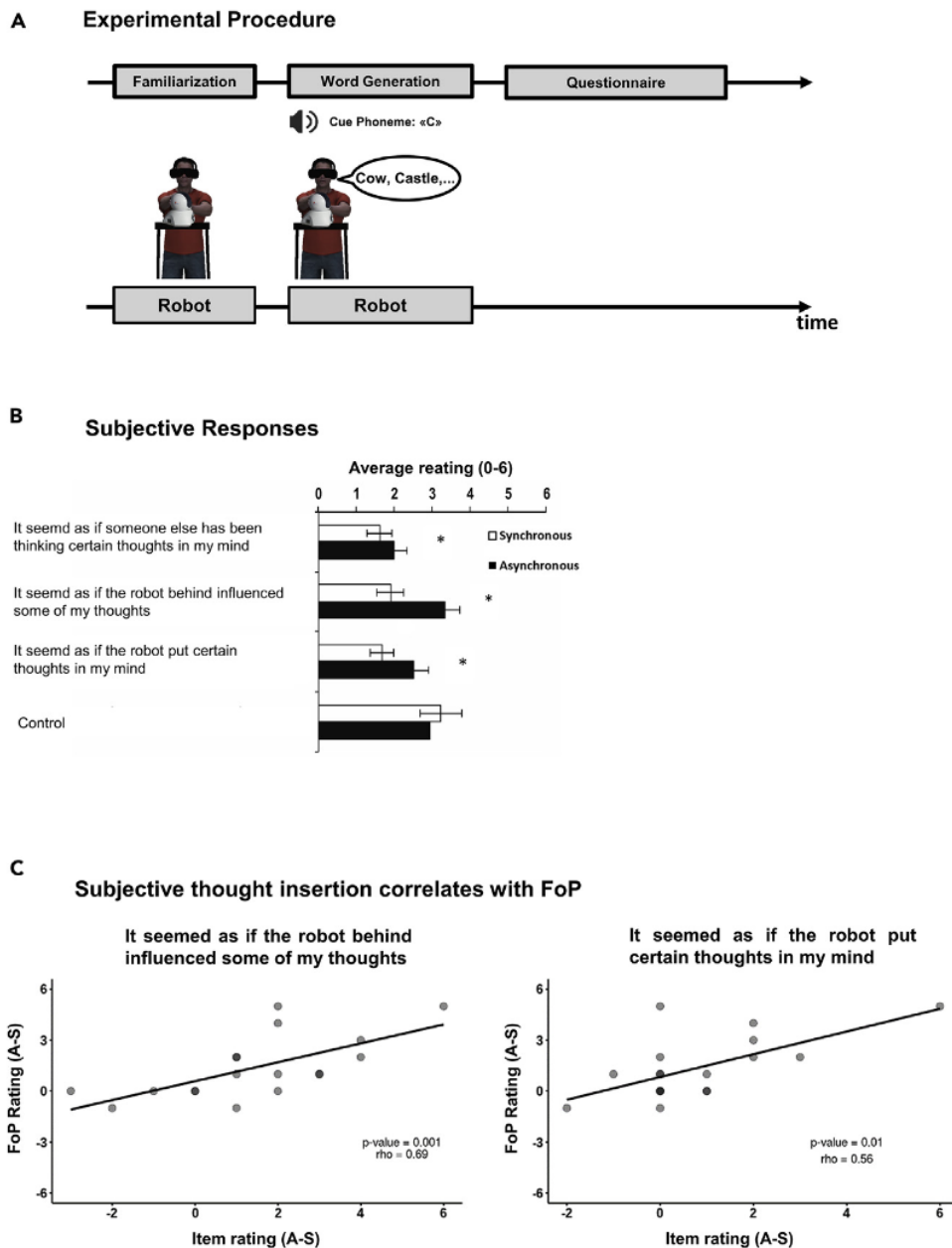


Figure 3. Thought insertion (experiment 3)

(A) At the beginning of each condition, participants heard a phoneme and then had 3 minutes to generate as many words starting with the specified phoneme as they could (Transparent methods).

(B) Subjective responses show that during the FoP-inducing asynchronous condition, participants agreed more with statements about TI and influencing.

(C) Correlation between FoP scores and thought-related experience ratings revealed a significant positive correlation between the differential FoP score and differential ratings of items reflecting TI and thought influencing. Error bars show standard errors of the mean. * $p < 0.05$.

Appendix

effect of suggestibility on the questionnaire items and further highlight the selectivity of the effects of sensorimotor stimulation and associated FoP on thought experience.

To summarize, the results from Experiment 3 demonstrate that repetitive spatiotemporal sensorimotor conflicts, while performing a verbal fluency task, induce sensations of thought alienation in healthy subjects. These sensations are weaker in intensity, but mimic aspects of the phenomenology of TI and thought influence as reported by psychiatric patients with delusions. We again induced the FoP in the same (asynchronous) experimental condition and we, importantly, show that the stronger our participants felt to be in the presence of an alien agent (FoP), the stronger they felt that somebody else was thinking or influencing thoughts in their mind, showing that subjective and behavioral TI can be induced and modulated experimentally using sensorimotor stimulation during a repetitive verbal fluency task (Experiments 2 and 3). More work is needed to follow-up on the results of Experiment 3. Thus, two main TI items ("It seemed as if someone else has been thinking certain thoughts in my mind"; "It seemed as if the robot put certain thoughts in my mind") showed higher ratings in the asynchronous FoP-inducing condition, whereas this was not the case for another TI item ("It seems as if some outside force or person is putting thoughts into my mind"). Future work should determine key phenomenological characteristics of TI in healthy participants when exposed to the present robotic system, focusing on the source of inserted thoughts and how thought ownership and thought agency are involved. This work should also determine how subjective aspects of TI potentially differ among individuals along the schizophrenia spectrum, how subjective TI relates to the implicit behavioral changes we observed, and how this depends on the involved cognitive task and sensorimotor stimulation.

Robotic-induced differences in thought-related source monitoring does not depend on differences in attentional demands

Results from Experiment 1 and Experiment 2 showed that the induced differences in self-monitoring during word memory and thought numerosity were specific for the asynchronous condition, were related to the experience of the alien agent (FoP), and did not manifest as a generic decrease in tasks performance; they were characterized by a specific reduction thought-related source monitoring (difference between active/self and passive/other processes). However, it could be argued that the higher level of sensorimotor incongruency in FoP-inducing asynchronous stimulation condition (compared with the synchronous condition) may have caused the described differences. Such an additional factor may have distracted participants, in turn more strongly affecting their SGE and thought numerosity judgments. To exclude this possibility, we tested the effects of robotic stimulation in the synchronous and asynchronous condition on a classic working memory 2-back task, chosen to tap into different mechanisms than source monitoring, while being well-known to require high-level attentional resources. If the effects of asynchronous stimulation depend on differences in attentional load between both conditions, then a reduction of working memory performance is expected specifically in the asynchronous condition. Conversely, the absence of a performance difference would rule out an attentional account, further corroborating our previous control analyses and supporting the conclusion that the robotic stimulation specifically affects source monitoring processes for internal thoughts, and not generically any cognitive process.

Appendix

As expected, at the subjective level, questionnaire responses showed that participants reported higher scores in the questions assessing the FoP (“I felt as if someone was standing behind my body”) ($Z = 20$, $p < 0.03$, one-tailed; Wilcoxon) and passivity experiences (“I felt as if someone else was touching my body”; $Z = 12$; $p < 0.01$, one-tailed; Wilcoxon). However, the pattern of stimulation did not affect the performance in the working memory task, as there was no difference between conditions in task accuracy ($t(1,19) = 0.26$, $p = 0.54$; Cohen’s $d = -0.14$; synchronous condition, mean accuracy = 92.1%; SD = 5.4; asynchronous condition: mean = 91.7; SD = 5.8). Differently from the previous tasks aimed at measuring the effects of the robot on internal thought processes—i.e., the SGE, Experiment 1, and the thoughts numerosity task, Experiment 2—the performance in the working memory (WM) task was unrelated to the FoP effect. Indeed, when we added the FoP score (i.e., the asynchronous-synchronous difference in the FoP questionnaire) as a covariate, we did not find any difference in performance between conditions ($F(1,19) = 1.83$, $p = 0.19$, $h^2 = 0.86$), or any interaction with the FoP score ($F(1,19) = 0.63$, $p = 0.44$, $h^2 = 0.29$). Thus, the robotic sensorimotor stimulation did induce a FoP in the asynchronous condition during a working memory task, but this did not alter participants’ performance in such a demanding cognitive task. To provide further support to this conclusion, we also run Bayesian statistics allowing us to measure how confidently we can accept the null hypothesis of no difference between conditions. The Bayesian factor was 0.41 (error 0.0002), suggesting a moderate evidence for the null hypothesis. Data from Experiment 4, therefore, suggest that asynchronous sensorimotor stimulation and related FoP do not induce a generic reduction of attentional resources affecting cognitive performance in general, supporting the conclusions from Experiments 1–3 about a specific effect on source monitoring of one’s own internal thoughts.

Discussion

Taken together, the behavioral data from Experiments 1–4 show that sensorimotor conflicts, applied during mental operations, reliably induce behavioral changes in thought-related source monitoring (SGE, perceived word numerosity), accompanied by alterations in thought consciousness that are compatible with some aspects of TI that are usually only seen in clinical populations. Importantly, these behavioral changes in conditions with increased TI are characterized by a reduced ability to discriminate mental processes representing self-generated thoughts from those generated by others, reducing the SGE for self-generated versus other-generated words (Experiment 1) and making thought numerosity judgments more similar for words that were either actively generated or passively heard (and generated by another person) (Experiment 3). These effects were especially observed in individuals experiencing an alien presence that we induced by asynchronous sensorimotor stimulation, showing that our robotic manipulation of thought-related source monitoring is not just associated with the loss of thought agency and TI but also with the feeling of the presence of an alien agent. Control analyses and the data from the control Experiment 4 further show that these effects cannot be explained by general differences in cognitive load between the two sensorimotor conditions.

Previous work has shown that the robot-induced FoP results from the manipulation of sensory and motor stimuli, which involve tactile stimulation on the back, as well as proprioceptive, tactile, and motor cues (from the upper limb), the congruency of which in the spatial and temporal domains are controlled via the robot. In the current and our previous research (Blanke et al., 2014; Salomon et al.,

Appendix

2020; Bernasconi et al., 2020), two main experimental conditions were used (synchronous and asynchronous sensorimotor stimulation). Both conditions contain a spatial conflict (between the spatial position of the moving hand and the spatial position of the touch cue delivered on the back of the participants), whereas the asynchronous condition also contains an additional spatiotemporal conflict (i.e., movement performed by the hand is delivered to the back of the participants with a delay of 500 ms). The present FoP setup was motivated by models of sensorimotor processing and the forward model of motor control (Wolpert and Ghahramani, 2020) that have been applied to bodily illusions and hallucinations (Fletcher and Frith, 2009). Previous reports have tested the effects of systematically varied sensorimotor conflicts (i.e., delays) on different hand-related bodily sensations (Weiskrantz et al., 1971; Blakemore et al., 1998, 2000) and the sense of agency (i.e., Farrer and Frith, 2002). However, there is an important additional element, compared with this previous research that is crucial for FoP induction: the interference with full-body processing (feedback at the back) that represents the body of the subject more globally and is a distinct sensorimotor and multi-sensory cortical system, when compared with the more local hand-related body representation system that has been studied by most previous investigators (for reviews see Blanke and Metzinger, 2009; Blanke et al., 2015). Thus, based on electrically induced FoP in a neurological patient (Arzy et al., 2006) and previous data using the same robotic system in healthy participants, Blanke et al. (2014) proposed that the FoP results from conflicting sensorimotor full-body signals that lead to the generation of a second self-representation (see self-location data in Blanke et al., 2014, and clinical data in Arzy et al., 2006) behind the participant that is misperceived as another person. Importantly, interference with this full-body system has been shown not only to lead to FoP but also to be related to global body illusions, such as out-of-body illusions and full-body illusions (i.e., Ehrsson, 2007; Lenggenhager et al., 2007).

The present data on TI demonstrate that the same asynchronous sensorimotor stimulation, when applied during mental operations, induces behavioral changes in thought-related source monitoring and in thought consciousness that depend on the FoP. For example, in Experiment 2, we showed that the intensity of the experimentally induced FoP was associated with changes in source monitoring characterized by self-generated words being perceived as more similar to other-generated words in the FoP-inducing asynchronous condition. As the number of self-generated words was perceived as higher and more similar to the number of other-generated words, selectively in the FoP-inducing asynchronous condition, we argue that under these conditions additional and alien-like thoughts were inserted into the minds of our healthy participants. In other words, if participants, while experiencing the FoP, are concurrently engaged in a cognitive task that implies implicit monitoring about the source of internal thoughts, the second self-representation behind the participant that is misperceived as another person (FoP), impacts such cognitive operations by inducing a misattribution of own inner thoughts. Accordingly, we argue that the present conflicting asynchronous sensorimotor stimulation in active, self-generating, conditions induces, in those participants experiencing the FoP, a mental state that is comparable to (albeit to a lesser degree and of short duration) to TI and thought alienation that is usually only reported by psychotic patients.

By defining a novel procedure that links robotics and cognitive science for the investigation of thought consciousness and its aberrations, the present approach offers, in healthy participants, novel insights into an enigmatic and clinically relevant psychotic symptom by firmly linking it to source monitoring and the FoP. Abnormal source monitoring has been shown to elegantly explain certain psychotic bodily experiences (i.e., somatic passivity, Frith, 1987) and has been proposed to account for other first-rank symptoms (Feinberg, 1978; Frith, 1987; Ford and Mathalon, 2004; Shergill et al., 2005) (i.e.,

Appendix

delusions of control; auditory verbal hallucinations), but had only limited success in explaining TI (Frith, 1987, 2005). Importantly, previous research was not able to manipulate TI experimentally and especially not able to induce TI-related mental states repeatedly and in controlled fashion (e.g., based on reaction time or accuracy measures) (Walsh et al., 2015; Sugimori et al., 2011). Central to our report is the experimental induction and manipulation of behavioral and subjective aspects of TI in healthy subjects, providing implicit-behavioural (SGE; thought numerosity) and explicit-subjective (questionnaires) data that conflicting sensorimotor stimulation is sufficient to induce alterations in thought consciousness when participants perform active mental operations. Behaviorally, we demonstrate that the present robotically induced TI is characterized by reduced source monitoring, a reduced ability to discriminate mental processes representing one's own mental operations from those representing mental operations of others, resembling passive thoughts and thoughts generated by another person, rather than one's own thoughts. Importantly, by manipulating specific sensorimotor processes that alter body representation (Blanke et al., 2014), we show that these changes were especially prominent in individuals experiencing an illusory and experimentally induced alien presence, as if the illusory alien presence (FoP) inserted alien thoughts into the mind of our healthy participants. We conclude that the present asynchronous sensorimotor stimulation induces in healthy participants, who tend to experience the illusory FoP, a mild and short-lasting behavioral and mental state that is reminiscent of symptomatic TI, an enigmatic and key symptom in psychosis.

Limitations of the study

The current study has some limitations. First, although the behavioral responses related to TI we report are robust and based on many repeated trials, the induction of subjective TI in study 3 was of mild to moderate intensity and thus differs from more prolonged and intense symptomatic TI in psychotic patients. Future work should strive to induce mental states of subjective TI of stronger intensity and also test the present robotic procedure and paradigms in psychotic patients with symptomatic TI. Second, in the present study only one condition of asynchronous stimulation was tested to induce the FoP. Thus, no specific indications can be provided about the critical delay between movement and feedback generating the effect. Other data available as pre-print (Bernasconi et al., 2020) show that the intensity of FoP rises depending on the delay between the moving hand and tactile feedback on the back (i.e., from 0 to 500 ms), with a plateau at 500-ms delay. Third, we did not investigate the involved neural correlates of TI or the FoP, which was outside the scope of the present study. Future work should target FoP and TI, jointly with brain imaging methods, in healthy participants and different patient populations, to unravel the brain mechanisms of robot-induced TI and FoP.

Finally, the present data do not allow us to indicate why some healthy participants are more prone to experience the FoP via our robotic sensorimotor stimulation. Individual differences in proneness to perceptual illusions (see, i.e., Marotta et al., 2016) and to bodily illusions (e.g., the rubber hand illusion; Asai et al., 2011; Tsakiris et al., 2011) have been extensively reported. However, different explanations have been proposed and will also apply to FoP. These range from differences in personality traits (e.g., hypnotizability: Lush et al., 2020; schizotypal and empathic traits: Tsakiris et al., 2011; sensory suggestibility: Marotta et al., 2016; perceptual priors within a Bayesian framework: Tulver et al., 2019) to neural differences such as differences in gray matter and in structural and functional connectivity (i.e., Kanai and Rees, 2011). It is possible that these non-mutually exclusive

Appendix

factors also contribute to individual differences in susceptibility to the FoP, but future research is needed to identify their specific roles.

Resource availability

Lead contact: Olaf Blanke. Center for Neuroprosthetics, School of Life Sciences, Campus Biotech Swiss Federal Institute of Technology. E-mail: olaf.blanke@epfl.ch.

Materials availability

This study did not generate new unique materials.

Data and code availability

The datasets generated during this study are available at Serino, Andrea (2020), “Thought consciousness and source monitoring depend on robotically-controlled sensorimotor conflicts and illusory states,” Mendeley Data, V1, <https://doi.org/10.17632/n2k4tjxzg8.1>.

METHODS

All methods can be found in the accompanying Transparent methods supplemental file.

SUPPLEMENTAL INFORMATION

Supplemental Information can be found online at <https://doi.org/10.1016/j.isci.2020.101955>.

ACKNOWLEDGMENTS

We thank Nathan Faivre and Jevita Potheegadoo for input on the manuscript, Elisa Ciaramelli for her important input on the SGE paradigm, and all members of the Blanke laboratory for discussions and other help with the project. This work was supported by two generous donors advised by CARIGEST SA (the first one wishing to remain anonymous and the second one being Fondazione Teofilo Rossi di Montelera e di Premuda), the Roger de Spoelberch Foundation, the Bertarelli Family Foundation, and the Swiss National Science Foundation to O.B. as well as by the National Center of Competence in Research: SYNAPSY –The Synaptic Bases of Mental Disease (financed by the Swiss National Science Foundation) to O.B. M.H. was supported by Grants-in-Aid for Scientific Research (B) of the Japan Society for the Promotion of Science.

Appendix

AUTHORS CONTRIBUTION

A.S., F.B., P.P. designed the study, carried out the experiments, analyzed data, and wrote the paper, M.S. carried out the experiments and analyzed data, M.H. and H.B designed and built the robotic device, P.P., K.D., J.P., and P.C. carried out clinical work, M.M., G.S., and H.D. carried out the experiments, A.G. and R.S. analyzed data, G.R. designed the study, built the robotic device, collected data, analyzed data, and wrote the paper; O.B. designed the study, analyzed the data, and wrote the paper.

DECLARATION OF INTERESTS

O.B. and G.R. are inventors of a granted US patent 10,349,899 B2 (System and method for predicting hallucinations, 2019). O.B., G.R., and M.H. are inventors of a granted US patent 10,286,555 B2 (Robot-controlled induction of the feeling of a presence, 2019). O.B. and G.R. are founders, shareholders, and members of the board of directors of Metaphysiks Engineering SA (Switzerland). O.B. is member of the board of directors of Mindmaze SA (Switzerland). The other authors do not have any competing interests to declare.

Received: August 14, 2020

Revised: October 27, 2020

Accepted: December 14, 2020

Published: January 22, 2021

REFERENCES

Arzy, S., Seeck, M., Ortigue, S., Spinelli, L., and Blanke, O. (2006). Induction of an illusory shadow person. *Nature* 443, 287.

Asai, T., Mao, Z., Sugimori, E., and Tanno, Y. (2011). Rubber hand illusion, empathy, and schizotypal experiences in terms of self-other representations. *Conscious. Cogn.* 20, 1744–1750.

Bays, P.M., Flanagan, J.R., and Wolpert, D.M. (2006). Attenuation of self-generated tactile sensations is predictive, not postdictive. *PLoS Biol.* 4, e28.

Bays, P.M., Wolpert, D.M., and Flanagan, J.R. (2005). Perception of the consequences of self- action is temporally tuned and event driven. *Curr. Biol.* 15, 1125–1128.

Appendix

- Bernasconi, F., Blondiaux, E., Potheegadoo, J., Stripeikyte, G., Pagonabarraga, J., Bejr-Kasem, H., Bassolino, M., Akselrod, M., Martinez-Horta, S., Sampedro, F., et al. (2020). Neuroscience robotics reveals fronto-temporal brain mechanisms of sensorimotor hallucinations in patients with Parkinson's disease. *Sci. Transl. Med.* <https://doi.org/10.1101/2020.05.11.054619>.
- Blakemore, S.J., Wolpert, D.M., and Frith, C.D. (1998). Central cancellation of self-produced tickle sensation. *Nat. Neurosci.* 1, 635–640.
- Blakemore, S.J., Wolpert, D.M., and Frith, C.D. (2000). Why can't you tickle yourself? *Neuroreport* 11, R11–R16.
- Blanke, O., Pozeg, P., Hara, M., Heydrich, L., Serino, A., Yamamoto, A., Higuchi, T., Salomon, R., Seeck, M., Landis, T., et al. (2014). Neurological and robot-controlled induction of an apparition. *Curr. Biol.* 24, 2681–2686.
- Blanke, O., and Metzinger, T. (2009). Full-body illusions and minimal phenomenal selfhood. *Trends Cogn. Sci.* 13, 7–13.
- Blanke, O., Slater, M., and Serino, A. (2015). Behavioral, neural, and computational principles of bodily self-consciousness. *Neuron* 88, 145–166.
- Ehrsson, H.H. (2007). The experimental induction of out-of-body experiences. *Science* 317, 1048.
- Faivre, N., Vuillaume, L., Bernasconi, F., Salomon, R., Blanke, O., and Cleeremans, A. (2020). Sensorimotor conflicts alter metacognitive and action monitoring. *Cortex* 124, 224–234.
- Farrer, C., and Frith, C.D. (2002). Experiencing oneself vs another person as being the cause of an action: the neural correlates of the experience of agency. *Neuroimage* 15, 596–603.
- Feinberg, I. (1978). Efference copy and corollary discharge: implications for thinking and its disorders. *Schizophrenia Bull.* 4, 636.
- Fletcher, P.C., and Frith, C.D. (2009). Perceiving is believing: a Bayesian approach to explaining the positive symptoms of schizophrenia. *Nat. Rev. Neurosci.* 10, 48–58.
- Ford, J.M., and Mathalon, D.H. (2004). Electrophysiological evidence of corollary discharge dysfunction in schizophrenia during talking and thinking. *J. Psychiatr. Res.* 38, 37–46.
- Frith, C.D. (1987). The positive and negative symptoms of schizophrenia reflect impairments in the perception and initiation of action. *Psychol. Med.* 17, 631–648.
- Frith, C.D. (2005). The self in action: lessons from delusions of control. *Conscious. Cogn.* 14, 752–770.
- Gallagher, S. (2004a). Agency, ownership, and alien control in schizophrenia. *Adv. Consciousness* 59, 89–104.

Appendix

Gallagher, S. (2004b). Neurocognitive models of schizophrenia: a neurophenomenological critique. *Psychopathology* 37, 8–19.

Hara, M., Rognini, G., Evans, N., Blanke, O., Yamamoto, A., Bleuler, H., and Higuchi, T. (2011). A novel approach to the manipulation of body- parts ownership using a bilateral master-slave system. *IEEE/RSJ Int. Conf. Intell. Robots Syst.* 4664–4669.

Hoffman, R.E. (1986). Verbal hallucinations and language production processes in schizophrenia. *Behav. Brain Sci.* 9, 503–517.

Johns, L.C., Cannon, M., Singleton, N., Murray, R.M., Farrell, M., Brugha, T., Bebbington, P., Jenkins, R., and Meltzer, H. (2004). Prevalence and correlates of self-reported psychotic symptoms in the British population. *Br. J. Psychiatry* 185, 298–305.

Kanai, R., and Rees, G. (2011). The structural basis of inter-individual differences in human behaviour and cognition. *Nat. Rev. Neurosci.* 12, 231–242.

Koehler, K. (1979). First rank symptoms of schizophrenia: questions concerning clinical boundaries. *Br. J. Psychiatry* 134, 236–248.

Lenggenhager, B., Tadi, T., Metzinger, T., and Blanke, O. (2007). Video ergo sum: manipulating bodily self-consciousness. *Science* 317, 1096–1099.

Lezak, M., Howieson, D., and Loring, D. (1995). Executive functions and motor performance. *Neuropsychol. Assess.* 3, 650–685.

Lush, P., Botan, V., Scott, R.B., Seth, A.K., Ward, J., and Dienes, Z. (2020). Trait phenomenological control predicts experience of mirror synaesthesia and the rubber hand illusion. *Nat. Commun.* 11, 1–10.

Marotta, A., Tinazzi, M., Cavedini, C., Zampini, M., and Fiorio, M. (2016). Individual differences in the rubber hand illusion are related to sensory suggestibility. *PLoS One* 11, e0168489.

Martin, J.R., and Pacherie, E. (2013). Out of nowhere: thought insertion, ownership and context-integration. *Conscious. Cogn.* 22, 111–122.

Miller, T.J., McGlashan, T.H., Woods, S.W., Stein, K., Driesen, N., Corcoran, C.M., Hoffman, R., and Davidson, L. (1999). Symptom assessment in schizophrenic prodromal states. *Psychiatr. Q.* 70, 273–287.

Orepic, P., Rognini, G., Kannape, O.A., Faivre, N., and Blanke, O. (2020). Sensorimotor Conflicts Induce Somatic Passivity and Loud Quiet Voices in Healthy Listeners. <https://doi.org/10.1101/2020.03.26.005843>.

Appendix

Salomon, R., Progin, P., Griffa, A., Rognini, G., Do, K.Q., Conus, P., Marchesotti, S., Bernasconi, F., Hagmann, P., Serino, A., and Blanke, O. (2020). Sensorimotor induction of auditory misattribution in early psychosis. *Schizophr Bull.* 8, 947–954.

Schneider, K. (1959). *Clinical Psychopathology* (Grune and Stratton).

Schultze-Lutter, F., Addington, J., Ruhrmann, S., and Klosterkötter, J. (2007). Schizophrenia Proneness Instrument, Adult Version (SPI-A) (Giovanni Fioriti).

Shergill, S.S., Bays, P.M., Frith, C.D., and Wolpert, D.M. (2003). Two eyes for an eye: the neuroscience of force escalation. *Science* 301, 187.

Shergill, S.S., Samson, G., Bays, P.M., Frith, C.D., and Wolpert, D.M. (2005). Evidence for sensory prediction deficits in schizophrenia. *Am. J. Psychiatry* 162, 2384–2386.

Shergill, S.S., White, T.P., Joyce, D.W., Bays, P.M., Wolpert, D.M., and Frith, C.D. (2014). Functional magnetic resonance imaging of impaired sensory prediction in schizophrenia. *JAMA Psychiatry* 71, 28–35.

Slamecka, N.J., and Graf, P. (1978). The generation effect: delineation of a phenomenon. *J. Exp. Psychol. Hum. Learn. Mem.* 4, 592.

Stephens, G.L., and Graham, G. (2000). *When Self-Consciousness Breaks: Alien Voices and Inserted Thoughts* (The MIT press).

Sugimori, E., Asai, T., and Tanno, Y. (2011). Sense of agency over thought: external misattribution of thought in a memory task and proneness to auditory hallucination. *Conscious. Cogn.* 20, 688–695.

Tsakiris, M., Jimenez, A.T., and Costantini, M. (2011). Just a heartbeat away from one's body: interoceptive sensitivity predicts malleability of body-representations. *Proc. R. Soc. B Biol. Sci.* 278, 2470–2476.

Tulver, K., Aru, J., Rutiku, R., and Bachmann, T. (2019). Individual differences in the effects of priors on perception: a multi-paradigm approach. *Cognition* 187, 167–177.

Vicente, A. (2014). The comparator account on thought insertion, alien voices and inner speech: some open questions. *Phenomenol. Cogn. Sci.* 13, 335–353.

Walsh, E., Oakley, D.A., Halligan, P.W., Mehta, M.A., and Deeley, Q. (2015). The functional anatomy and connectivity of thought insertion and alien control of movement. *Cortex* 64, 380–393.

Weiskrantz, L., Elliott, J., and Darlington, C. (1971). Preliminary observations on tickling oneself. *Nature* 230, 598–599.

Wolpert, D.M., and Ghahramani, Z. (2020). Computational principles of movement neuroscience. *Nat. Neurosci.* 3, 1212–1217.

Appendix

Transparent Methods

Participants

A total of 93 healthy participants took part in four separate behavioural experiments. Experiment 1 consisted of 35 participants (11 female; mean age: $M=20.5$ years, $SD=2.5$ years), Experiment 2 of 19 participants (9 female; mean age: $M=20.3$ years, $SD=2.4$ years), and Experiment 3 of 19 participants (6 female; mean age $M=20.9$ years, $SD=2.0$ years), and Experiment 4 of 20 participants (10 female; mean age $M = 28.4$ years, $SD = 6.29$ years). All participants for Experiments were recruited by an advertisement at the EPFL campus (École Polytechnique Fédérale de Lausanne, Lausanne, Switzerland) and at Campus Biotech (Geneva, Switzerland). All participants had normal touch perception and no psychiatric or neurologic history as assessed by self-report. All participants were native French speakers. Each participant only took part in one experiment only. All participants were naive to the purpose of the experiments and gave written informed consent to take part in the experiment. Experiments were approved by the EPFL ethics committee (Comité d'éthique de la recherche humaine) and were conducted according to the ethical standards laid down in the Declaration of Helsinki. Participants gave written informed consent after the experimental procedures were explained to them and were reimbursed for their participation with 20 Swiss Francs.

Apparatus

Robotic sensorimotor system. To experimentally create sensorimotor mismatch we adapted a bilateral master-slave robotic system that has been previously used to manipulate changes in bodily self-consciousness (Blanke et al., 2014; Hara et al., 2011). This system is composed of a commercial master haptic interface, the Phantom Omni (Sens Able Technologies), and a three degree-of-freedom slave robot. The slave device consists of two mechanisms: a belt-drive mechanism and a parallel-link mechanism. The belt-drive mechanism is made up of a belt linked to a direct-drive DC motor (RE 40, Maxon) moving a carrier on a linear guide allowing movements in the y (forward-backward) direction. The parallel-link mechanism is actuated through two harmonic drive motors (RH-8D 6006, Harmonic Drive Systems) and enables both tapping and stroking in x (right-left) and z (up-down) directions. These three motors equipped with optical encoders for positions sensing are connected to motor drivers (4-Q-DC Servoamplifier LSC 30/2 & ADS 50/5, Maxon) that receive the command voltages from a computer via PCI data acquisition cards (NI PCI-6221 & NI PCI-6014, National Instruments). The overall workspace of the slave device is 200mm in the x direction, 250mm in the y direction, and 200mm in the z direction (See Figure S1).

A load cell (ELPFTIM-50N, Measurement Specialties) is attached to the tip of the slave device in order to measure contact force. This allowed us to introduce a compliance factor on the system preventing the slave device from applying instantaneous strong force to the participants, making the interaction safer and more realistic. The system was controlled through an application programmed in Visual C++ (Microsoft) at a sampling rate of 1 kHz. The latency related to information transfer delays and computational processing necessary for mapping the master device movements to the slave device movements (i.e. touching the back of the participants) was equal to 1ms. The system had a bandwidth of approximately 2.5 Hz allowing a good synchrony between the master and the slave even during rapid and abrupt changes in velocity and direction (Hara et al., 2011). This allowed reducing the constraints on participants' movements.

Appendix

In each experiment, the participants were first explained the task and informed about the general procedure of the experiment. Then they were instructed on how to use the robotic device to apply touch on their back through the tip of the slave device. The experimenter demonstrated the type of movements they were supposed to perform during the experimental blocks. In particular, they were asked to perform tapping movements in front of them by holding the master device with both hands, while receiving the touch on their back by the slave device. They were allowed to tap in different directions (up-down, left-right) resulting in different touches applied on their back within a workspace of 200x200mm. In the training session, the participants used the system in the synchronous mode for about 1 minute without being blindfolded.

General experimental procedure

The robotic sensorimotor system was used to apply sensorimotor stimulation in the different experiments in two different conditions: synchronous sensorimotor stimulation (the participants were asked to move the lead robot via their right index finger, this way actuating the follow robot which provided immediate and congruent touches to the participant's back) and asynchronous sensorimotor stimulation (500 ms delay between the first robot operated via the right index finger and the second robot applying tactile feedback on the participants' back). During the robotic stimulation participants were always blindfolded. In each experiment, the participants were first explained the task and informed about the general procedure of the experiment. Then they were instructed on how to use the robotic device to apply touch on their back through the tip of the follow device. The experimenter demonstrated the type of movements they were supposed to perform during the experimental blocks. In particular, they were asked to perform tapping movements in front of them by holding the lead device, while receiving the touch on their back by the follow device. They were allowed to tap in different directions (up-down, left-right) resulting in different touches applied on their back within a workspace of 200x200mm. In the training session, the participants used the system in the synchronous mode for about 1 minute without being blindfolded.

Experiment 1 – self generation effect: design, procedure and analyses

Experiment 1 was designed to assess the so-called self-generation effect (SGE, Slamecka et al., 1978), originally described by Slamecka and Graf, while performing robotic stimulation. In this paradigm, the participants are either presented with a list of words (passive condition; participants only heard the words) or they were asked to generate their own words (active condition; participants produced and heard the words) within a given set of rules (see next two paragraphs for more detail). During the encoding session, we asked participants to memorize both the heard and the self-generated words. In the recognition session, participants were presented with a list of words (pre-recorded and played back), containing either the words they had generated or heard and other semantically related words, that were never presented and used as distractor (50% of target and 50% of distractor words were presented in random order). Participants were asked to determine for each word whether it is a word he or she had generated or heard during the encoding session or not. Participants typically remember the self-generated items (active condition) better than the heard items (passive condition). This phenomenon, termed self-generated effect, SGE, has been shown to be very robust and has been

Appendix

described in recognition and recall tasks and with a variety of materials, generation rules, and retention intervals (Hirshman et al., 1988).

Active condition. In the active conditions, participants heard 35 cue words, each followed by a cue letter. Participants were instructed to generate an associated word, which had to start with the specified letter, and utter it out loud. If the participant's generated word matched the predicted word (target word), the experimenter registered it, and the word was later used in the recognition task during the test recognition phase. The time interval between the cue word and cue letter presentation was 1s, and participants' performance was self-paced in the encoding as well as in the recognition session.

Passive Condition. In the encoding session of the passive conditions, participants merely listened to 35 audio-played word pairs, and were instructed that they would be later tested for recognition of the second word in a pair.

Design and procedure. We performed a 2 x 2 factorial repeated measures design with the factor Sensorimotor stimulation (synchronous and asynchronous sensorimotor stimulation) and the factor Source (active, passive). Each participant therefore completed four experimental conditions, given in randomized order. At the beginning of each condition and before the encoding session started, we asked participants to move the robot for 60 seconds. During the encoding session, depending on the condition, they listened to either pairs of words (passive conditions) or they generated their own words after hearing a cue word and a cue letter (active conditions), while continuing to operate the robotic device. Participants wore headphones and were blindfolded throughout the encoding phase. After the encoding session, participants were asked to stop operating the robot and to remove the blindfolding and commenced the recognition task. At the end of the task and after each condition, participants were administered the questionnaire (see below).

Effect size estimation. We estimated the effect size for the self-effect for Experiment 1 based on the Experiment reported in Blanke et al., 2014 (study 4), where by participants were asked to estimate how many people there were close to them. Participants reported on average 0.74 (SD=0.30) and 0.99 (SD=0.34) persons, resulting in an effect size of 0.73 resulting on a suggested sample size of $N = 28$. We initially recruited 35 participants and we only included in the present analysis those participants, who produced enough word associations and whose performance was above chance (Slamecka et al., 1978). This sample size is in line with the original report of the Generation effect by (Slamecka et al., 1978) ($N = 24$).

Data analysis. In order to avoid ceiling effects in the recognition task for self-generated words, only the data of participants who generated more than 50% of expected associations (at least 18 words) were included into analysis. Also, participants who performed below chance level in the recognition task were excluded from analysis, leaving data from 22 remaining participants for further analysis. Task performance was defined by d-prime scores, which were then analyzed with repeated measures ANOVA, with Source and Sensorimotor stimulation as the two within-subjects factors and the feeling of a presence (FoP) score as a covariate. Based on the recent finding that the FoP can be experimentally induced in healthy participants due to a specific spatial and temporal sensorimotor mismatch (Blanke et al., 2014), we calculated the FoP score by subtracting the ratings of the FoP questionnaire item in the asynchronous from the ratings in the synchronous condition. Thus, higher FoP scores indicate a stronger FoP illusion due to the robotically induced sensorimotor mismatch.

Appendix

Acquisition and preparation of auditory word stimuli. For Experiment 1, 250 word association pairs were first selected from the database of word association norms containing a collection of French words (Ferrand et al., 1998). In order to balance the strength of association between the cue word and its associated target word across conditions, we have recruited 10 native French speakers (2 females; 18 – 23 years, $M=20.1$, $SD=1.66$). They were given the selected 250 cue words and cue letters (first letter of the predefined target word) to generate associations. The strength of the association was defined as the frequency with which participants chose the target word. 70 association pairs with higher association strength (0.7 - 1) were then selected for the self-generated conditions to increase the probability of participant generating the target word. 70 association pairs with lower association strength (0.3 - 0.6) were used for the other-generated conditions. Another 140 words were selected from the database to be used as distractor words during recognition task. The association pairs were then sorted into 4 alternative word lists (2 for self-generated and 2 for other-generated conditions), each consisting of 35 word pairs, with balanced association strength. Similarly, the distractor words were divided into 4 lists, each containing 35 distractor words. We verified, using the multivariate analysis of variance (MANOVA), that there was no significant difference in terms of frequency of use (www.lexique.org) and word length between the alternative lists ($F(6, 544)=0.494$, $p=0.813$) or between the target and distractor words ($F(2, 271)=0.001$, $p=0.999$). The word set was then recorded by two male and two female native French speakers and registered in wav format with 11025 Hz sampling frequency. In both Experiment 1 and Experiment 2, as well as during the pilot experiment (see below, Supplementary Results), the auditory word stimuli were played to participant in a gender-matched voice. In Experiment 1, two gender-matched voices were alternating between the encoding and testing phase in a balanced manner throughout the experiment.

Experiment 2 - Thought numerosity: design, procedure and analyses

Experiment 2 was designed to estimate the number of thoughts in the participants' mind while performing robotic stimulation (in analogy to perceptual numerosity tasks; Krueger, 1972). To this aim, we implemented a fluency task, whereby, we asked participants to estimate the number of words that they have either generated themselves (active condition) or have listened to (passive condition), while operating the robotic sensorimotor system.

Design. We used a 2 x 2 factorial repeated-measures design, whereby we manipulated the Sensorimotor stimulation (synchronous and asynchronous sensorimotor stimulation) and the Source of the words to be estimated (active, passive). In the active conditions, a starting phoneme was played to participants through the headphones and they were instructed to generate as many words as possible starting with the specified phoneme, in a given time period (phonetic fluency task). This time period randomly varied between 15 and 30s, in order to avoid participants always producing and estimating a similar number of words. The experimenter counted and registered the words and, immediately afterwards, the participant had to estimate how many words she or he had generated. In the passive conditions the participants listened to a list of words, consisting of between 6 and 10 words (based on the number of words another group of participants generated in the active condition; see Pilot experiment in Supplementary Results). The number of words randomly varied throughout the trials. The words were played to participants with an inter-stimuli interval of 2.5s. All words and phoneme cues were presented to participants as auditory stimuli using MATLAB software (MathWorks, Inc.). In the passive condition, in order to prevent participants from counting the words,

Appendix

they were asked to determine whether each word they heard contains a phoneme, specified at the beginning of a trial. Each condition was repeated three times, and each repetition consisted of 4 trials, resulting in total of 12 numerosity judgments per condition. The order of repetitions of the different experimental conditions was counterbalanced across the participants. The dependent variable was the numerosity judgement accuracy, calculated by subtracting the actual number of played or produced words from the number of judged number of words. Prior to the beginning of the experimental session, participants went through a training session, comprising one repetition of each condition. Before or after the experiment in a counterbalanced manner, participants were asked to operate the robot for 60s in the synchronous and asynchronous mode (run in counterbalanced order), being blindfolded and instructed to only focus on their movements and tactile feedback. After the synchronous and asynchronous blocks, they were given the FoP illusion questionnaire in order to measure the degree of the illusion induced by the sensorimotor stimulation.

Effect size estimation. Data from Experiment 1 were used to estimate the minimum sample size for Experiment 2 (self-effect part). In the group who experienced the FoP, the self-effect for asynchronous and synchronous stimulation was 1.53 (SD=0.83) and 2.33 (SD=0.75), resulting in an effect size of 1.008 and suggesting a minimum sample size of 15. We recruited 19 participants. For the questionnaire part, we estimated the required sample to replicate the FoP effect based on Study 3 from Blanke et al., 2014. The average ratings for the FoP question were 4 (SD=1.9) and 2.14 (SD=1.65) in the asynchronous and synchronous conditions respectively, resulting in a size effect of 1.30 and suggesting a sample size of 15 participants. We tested 19 participants via questionnaires assessing subjective thought insertion.

Data analysis. Two trials from two participants were discarded from analysis, because they failed to generate any word within the given time limit. The differences between the numerosity judgment and actual number of words (judgment accuracy) were averaged within each condition for each participant and then analyzed with repeated measures ANOVA where Sensorimotor stimulation (synchronous and asynchronous sensorimotor stimulation) and Source (active, passive) were used as within-subject factors.

Experiment 3 – changes in thoughts subjective experience: design, procedure and analyses
Experiment 3 was designed to measure whether robotic sensorimotor stimulation induced explicit changes in the subjective experience associated to internal thoughts. To this aim, a new group of participants again operated the robotic lead-follow system (as in Experiment 1 and 2), while simultaneously performing a phonetic fluency task. In a repeated-measures design, we manipulated the factor Sensorimotor stimulation (synchronous vs asynchronous). Participants manipulated the robotic system in synchronous and asynchronous mode for 3 minutes. At the start of each condition, they heard a French phoneme through headphones, and were then given three minutes to generate as many words as possible that started with the specified phoneme. At the end of each condition, they were asked to answer several questions referring to their thinking process during task performance (see below). The order of synchronous and asynchronous conditions was counterbalanced across the subjects. Before or after the experiment in a counterbalanced manner, participants were also asked to operate the robot for 60s in both, synchronous and asynchronous modes while blindfolded and focused on their bodily sensations. After these synchronous and asynchronous blocks, they were given the FoP questionnaire (see below).

Appendix

In order to evaluate subjective experience during internal thoughts, we designed a detailed, 12 - item questionnaire. The items were constructed based on the literature on thought possession disorders (Miller et al., 1999; Schultze-Lutter et al., 2007) and particularly targeted feelings related to thought insertion (ex. "It seemed as if some outside force or person has put certain thoughts in my mind"), thought influence (ex. "It seemed as if some outside force or person has influenced some of my thoughts"), thought ownership (ex. "It seemed as if certain thoughts I had belonged to someone else") and thought withdrawal (ex. "It seems as if some of my thoughts have been removed from my mind"). Other items, which served as control for suggestibility, pertained to positive psychotic symptoms, but not to disorders of thought possession, i.e. parasite thoughts, thought echoing, and voice distortion. The participants were asked to rate how much they agreed with each questionnaire item on a 7-point Likert scale (0 = not at all, 3 = not certain, 6 = very strong) (see Table S2).

Experiment 4 - working memory task - design, procedure and analyses

Experiment 4 was designed to assess the effects of robotic stimulation on a working memory task. As in the previous experiments, blindfolded participants operated the robotic lead-follow system, while they were performing a 2-back verbal task, selected as a well-established paradigm and highly demanding in terms of cognitive resources. In a repeated-measures design, we manipulated Sensorimotor stimulation (synchronous vs asynchronous). Participants performed 16, 24 or 32 second blocks of sensorimotor stimulation with the robot either in the synchronous or asynchronous conditions, in counterbalanced order. During the stimulation, they were presented with a series of numbers (one every two seconds), which were administered via headphones. Subjects were required to respond (via button press) if the current number in the series was equal to the last but one heard in the series. Each condition consisted of 24 trials. At the end of the task, they were also presented with the BSC questionnaire assessing the FoP and related sensations (see below).

Effect size estimation. Effect size was calculated based on the results of Experiment 2. Given the obtained differences between self-generated and other-generated words and the associated standard deviations, resulting in a significant interaction between stimulation condition and agent ($p=0.015$), the necessary sample size to replicate the effect with a $p<0.05$ is $N=20$.

Subjective changes in BSC.

To measure changes in bodily self-consciousness as induced by the robotic sensorimotor stimulation, for all experiments, we administered the same questionnaire as used previously (4). The questionnaire consists of 8 items, referring to the feeling of presence ("I felt as if someone was standing behind my body"), sensation of passivity ("I felt as if someone else was touching my body"), and other bodily illusions (see Supplementary Information). Two items served as control items for suggestibility (i.e. "I felt as if I had no body" and "I felt as if I had more than one body"). Participants were asked to designate on a 7-point Likert scale, how strongly they felt the sensation described by each item (0 = not at all, 3 = not certain, 6 = very strong).

Statistical Analysis. To assess statistical differences induced by the different experimental conditions on the subjective experiences (thoughts and BSC questionnaires), a one-tailed Wilcoxon signed rank test was applied to each question independently to compare response for the synchronous and

Appendix

asynchronous stimulation condition. One-tailed was decided because of the strong hypothesis that our effects would be always bigger for the asynchronous sensorimotor stimulation.

Supplemental Results

Experiment 1

FoP Questionnaire. Participants experienced stronger sensation to be touched by another person in the asynchronous condition (synchronous: $M = 3.20$, $SD = 1.71$; asynchronous: $M = 3.84$, $SD = 1.78$; $Z = -2.399$, $p = 0.005$) and also reported a stronger feeling of a presence in the same, asynchronous condition (synchronous: $M = 2.80$, $SD = 1.55$; asynchronous: $M = 3.24$, $SD = 1.48$; $Z = -2.361$, $p = 0.005$). Conversely, the participants reported stronger illusory self-touch when they operated the robotic system in the synchronous condition (synchronous: $M = 3.18$, $SD = 1.50$; asynchronous: $M = 2.24$, $SD = 1.43$; $Z = -2.985$, $p = 0.002$). The ratings of the control items were low and not significantly affected by the sensorimotor stimulation ($M < 1.5$, $SD < 1.80$, all $p < 0.05$), except the item: “I felt as if I was behind my body” (synchronous: $M = 2.30$, $SD = 1.77$; asynchronous: $M = 1.82$, $SD = 1.47$; $Z = -2.623$, $p = 0.005$).

Subjective loss of thought agency (Questionnaire). The effect of sensorimotor stimulation modulation was also observed for the sense of agency over self-generated thoughts. The participants reported a reduced sense of agency (“It seemed as if I was not the one who generated the words”) for the words they generated in the asynchronous ($M = 2.21$, $SD = 1.728$) as compared to the synchronous condition ($M = 1.73$, $SD = 1.625$; $Z = -1.894$, $p = 0.029$). The sensorimotor stimulation did not modulate the experience of thought insertion proper (synchronous: $M = 2.97$, $SD = 1.610$, asynchronous: $M = 3.03$, $SD = 1.794$; $Z = -.340$, $p = 0.367$) or ownership for self-generated words (synchronous: $M = 1.55$, $SD = 1.348$, asynchronous: $M = 1.67$, $SD = 1.407$; $Z = -0.537$, $p = 0.296$), although in both cases ratings were higher in the asynchronous condition.

Pilot Experiment: Verification of the generation effect with auditory stimuli. To confirm that the generation effect can be also achieved by using selected word stimuli presented in the auditory modality, prior to the main experiment we conducted a pilot experiment without the robotic sensorimotor stimulation. 6 native French speaking participants (3 females, $M = 20.6$ years, $SD = 2.7$) were recruited to participate in this experiment. They completed two self-generated and two other-generated conditions in a randomized order. Two-tailed paired-sample t-test showed that the accuracy rate ($t(5) = 5.289$, $p = 0.003$), as well as sensitivity ($t(5) = 7.264$, $p = 0.001$), in the recognition task was significantly higher for the self-generated words, demonstrating that the generation effect was replicated with the selected auditory word material.

Behavioral paradigm: Self-generation effect. The analysis of performance in the memory task replicates the classical self-generation effect (SGE), as the main effect of source was significant (self: $M = 4.12$, $SD = 0.45$; other: $M = 2.28$, $SD = 0.65$; $F(1,20) = 180.86$, $p < 0.0001$). Importantly,

this self-effect was significantly modulated by the manipulation of the sensorimotor stimulation in relation to the experience of FoP (interaction between generation source, sensorimotor stimulation and covariate FoP score: $F(1,20) = 6.95$, $p = 0.016$). To investigate this interaction, we split the sample

Appendix

into two groups according to the experience of FoP (No-FoP group: FoP score ≤ 0 ; FoP group: FoP score > 0) and tested whether the two groups differed in the modulation of the self-effect due to sensorimotor mismatch. The mixed ANOVA on the strength of the self-effect (calculated as d' for recognition of self-generated – d' for other generated words) showed a significant interaction between sensorimotor stimulation and group ($F(1,20) = 7.217$, $p = 0.014$). Post-hoc comparisons further showed a significant decrease of the self-effect in the asynchronous condition, but only in the group, which experienced the FoP (synchronous: $M = 2.33$, $SD = 0.75$; asynchronous: $M = 1.53$, $SD = 0.83$; one-tailed t-test: $t(10) = 2.148$, $p = 0.029$; No-FoP group: synchronous: $M = 1.43$, $SD = 0.90$; asynchronous: $M = 2.07$, $SD = 1.04$, one-tailed t-test: $t(10) = 1.660$, $p = 0.064$).

Experiment 2

The number of generated words in the thought numerosity judgment. In Experiment 2 (Numerosity judgment), we used the same auditory verbal stimuli as in Experiment 1 (in total 420 French words and 22 phonemes). To verify that the found differences in numerosity judgment were not due to the differences in the number of generated words between the experimental conditions, we conducted a repeated-measures ANOVA with source and sensorimotor stimulation as within- subject factors. The analysis showed that the number of generated words did not differ between the active ($M = 7.95$, $SD = 2.02$) and passive conditions ($M = 8.11$, $SD = 0.33$; $F(1,18) = 0.115$, $p = 0.738$), neither it was modulated by the sensorimotor stimulation (synchronous: $M = 8.18$, $SD = 1.18$; asynchronous: $M = 7.88$, $SD = 0.89$; $F(1,18) = 3.079$, $p = 0.096$) or the interaction between the source and sensorimotor stimulation ($F(1,18) = 0.944$, $p = 0.344$).

Absolute accuracy in the thought numerosity judgment. To verify whether the difference in the numerosity judgments was not due to differences in cognitive load between the active and passive conditions or between synchronous and asynchronous conditions, we analyzed the absolute accuracy. This was defined as a percentage of trials when the numerosity judgment was correct within each experimental condition. The repeated measures ANOVA showed that the absolute accuracy was not affected by the Source ($F(1,18) = 0.833$, $p = 0.374$), Sensorimotor stimulation ($F(1,18) = 0.810$, $p = 0.380$) or their interaction ($F(1,18) = 1.118$, $p = 0.304$).

Experiment 3

Analyses of the questionnaire data revealed that the synchrony between participants' movements and received tactile feedback significantly modulated ratings of the questionnaire items related to thought insertion and thought influencing. In particular, as compared to the synchronous, the asynchronous mode of stimulation resulted in significantly higher ratings of the items assessing thought insertion: "It seemed as if someone else has been thinking certain thoughts in my mind" (synchronous: $M = 1.61$, $SD = 1.38$, asynchronous: $M = 2.00$, $SD = 1.41$; $Z = 2.111$, $p = 0.03$), thought influencing: "It seemed as if the robot behind influenced some of my thoughts" (synchronous: $M = 1.89$, $SD = 1.49$, asynchronous: $M = 3.33$, $SD = 1.64$; $Z = 2.345$, $p = 0.01$) and a significant higher ratings of the item assessing robotically-induced thought insertion: "It seemed as if the robot put certain thoughts in my mind" (synchronous: $M = 1.67$, $SD = 1.33$, asynchronous: $M = 2.5$, $SD = 1.76$; $Z = 1.911$,

Appendix

$p = 0.03$). The ratings of other questionnaire items were not significantly modulated by the sensorimotor mismatch (all $p > 0.05$). See Table S2.

Supplemental References

Blanke, O. Pozeg, P., Hara M., Heydrich, L., Serino, A., Yamamoto, A., Higuchi, T., Salomon, R., Seeck, M., Landis, T., Arzy, S., Herbelin, B., Bleuler, H., and Rognini, G. (2014). Neurological and robot-controlled induction of an apparition. *Current Biology* 24, 2681-2686.

Ferrand, L., and Alario, F.X. (1998). Normes d'associations verbales pour 366 noms d'objets concrets. *L'Année psychologique* 98, 659-709.

Hara, M., Rognini, G., Evans, N., Blanke, O., Yamamoto, A., Bleuler, H., Higuchi, T. (2011). A Novel Approach to the Manipulation of Body-Parts Ownership Using a Bilateral Master- Slave System. *IEEE/RSJ International Conference on Intelligent Robots and Systems*, 4664-4669.

Hirshman, E. and Bjork, R.A. (1988). The generation effect: Support for a two-factor theory.

Journal of Experimental Psychology: Learning, Memory, and Cognition 14, 484. Krueger, L.E. (1972). Perceived numerosity. *Perception and Psychophysics* 11, 5-9.

Miller, T.J., McGlashan, T.H., Woods, S.W., Stein, K., Driesen, N., Corcoran, C.M., Hoffman, R., and Davidson, L. (1999). Symptom assessment in schizophrenic prodromal states. *Psychiatric Quarterly* 70, 273-287.

Schultze-Lutter, F., Addington, J., Ruhrmann, S., and Klosterkötter, J. (2007). Schizophrenia proneness instrument, adult version (SPI-A) (Rome: Giovanni Fioriti).

Slamecka, N.J., and Graf, P. (1978). The generation effect: delineation of a phenomenon.

Journal of experimental Psychology: Human learning and Memory 4, 592.

Appendix

6.3 Presence Hallucination during locomotor activities in patients with Parkinson's Disease

Jevita Potheegadoo ^{1,2,*}, **Herberto Dhanis** ^{1,2}, Judit Horvath ^{3,4}, Pierre R. Burkhard ⁴,
Olaf Blanke ^{1,2,4}

Affiliations

¹ Center for Neuroprosthetics, Faculty of Life Sciences, Swiss Federal Institute of Technology (EPFL), 1202 Geneva, Switzerland.

² Brain Mind Institute, Faculty of Life Sciences, Swiss Federal Institute of Technology (EPFL), 1015 Lausanne, Switzerland.

³ La Tour Hospital, 1217 Meyrin, Geneva, Switzerland.

⁴ Department of Clinical Neurosciences, Geneva University Hospital, 1205 Geneva, Switzerland.

Corresponding author

Dre. Jevita Potheegadoo
Center for Neuroprosthetics & Brain Mind Institute
School of Life Sciences, Campus Biotech
Ecole Polytechnique Fédérale de Lausanne (EPFL)
Chemin des Mines, 9
1202 Geneva
Switzerland
E-mail: jevita.potheegadoo@epfl.ch
Tel: +41 (0)21 693 95 68

Keywords: Parkinson's disease psychosis; non-motor symptoms; illusions; sensorimotor processing, procedural activities.

Appendix

Appendix

The presence hallucination (PH) is the sensation that somebody is nearby when no one is actually there. Affecting ~64% of patients with Parkinson's disease (PD) and occurring early, PHs are clinically relevant for indicating potential negative clinical outcome (1–3). Recently, we have induced PHs in PD patients by generating sensorimotor conflicts while patients repeatedly actuated a robotic device placed in front of them (4). Patients with symptomatic PHs were more sensitive to such sensorimotor stimulation than those without. We also identified abnormal sensorimotor processes predictive for the occurrence of PHs.

Here, we describe the case of two patients with PD who reported PHs after PD onset only when they were involved in repetitive locomotor activities in daily life - clinical evidence in favour of the importance of sensorimotor signals in PHs in PD, compatible with repetitive robotic sensorimotor stimulation inducing PH in PD (4).

A 73-year old patient experienced PHs repeatedly, but only when walking outside his home. PHs are frequent and enduring, occurring several times per week: "I'm walking, then I feel that someone is just behind me, a bit on my right, wanting to overtake me, bending over my shoulder. This person seems to walk faster than me, just two steps behind. I stop quickly, and move sideways for this person to pass, but when I look behind, there's no one". The patient also described passage hallucinations (2) and somatosensory hallucinations, all occurring outside locomotion periods.

Another 79-year old patient experienced two distinct PHs, one positive (guardian angel) and one negative PH, occurring independently from each other several times per month, only during walking. The patient has the impression of someone behind her, close (1m), on the right side, next to her shoulder, as if someone is "catching up" with her. The positive PH lasts for a few seconds only, the negative PH somewhat longer. The patient would step sideways to let the "person" pass but there is no one. Passage hallucinations and visual illusions occur occasionally, always outside locomotion periods.

The repeated occurrence of PHs during locomotion in these two patients, and not in other situations, suggests that sensorimotor processing related to gait plays a functional role in PH. We argue that for some patients with PD, locomotion – a highly procedural motor activity associated with tactile and proprioceptive feedback – facilitates PHs by interfering with sensorimotor processes (4). A widespread network including cerebellum, cortical and sub-cortical structures is involved in locomotion and integration of related sensorimotor signals, for which the role of the basal ganglia has particularly been emphasised (5). Partly similar brain regions were recruited in robot-induced PHs (4), suggesting a neural overlap between gait control and PHs. We note that both patients did not suffer from dementia, nor delusions (Table 1). PHs were unrelated to recent changes in medication.

Future research in patients with PD investigating the impact of ongoing locomotor activity on altered sensory perceptions is needed and may reveal that PHs during locomotion is more frequent than commonly thought.

Appendix

TABLE 1 Socio-demographic information, clinical data and medications of the two patients with Parkinson's disease (PD)

	Case 1	Case 2	
Socio-demographic data			
Age (years)	73	79	
Sex	M	F	
Level of schooling (years)	9	12	
Handedness	Right	Right	
Clinical characteristics of PD			
Duration of PD (years)	14	7	
Symptoms lateralization (onset)	Right	Left	
UPDRS 1	16/52	15/52	
UPDRS 2	19/52	15/52	
UPDRS 3	19/124	18/124	
UPDRS 4	1/24	0/24	
Medications			
Levodopa equivalent daily dose (including dopamine agonists) (mg/day)	1180	1900	
Neuropsychological tests			Cut-off scores
MoCA	26	27	< 26/30
PD-CRS (total score)	73a	102	≤ 81/134
Fronto-subcortical score (executive functioning)	43	72	
Posterior cortical score	30	30	
Occurrence of hallucinations			
Presence hallucinations (PH)	✓	✓	
Passage hallucinations	✓	✓	
Visual illusions (eg, pareidolia)	✓	✓	
Visual hallucinations	×	×	
Auditory hallucinations	×	✓	
Somatosensory hallucinations	✓	×	
Olfactory/gustatory hallucinations	×	×	

^aScores below normal range.

✓ presence of hallucination.

× absence of hallucination.

UPDRS, Unified Parkinson's disease rating scale; MoCA, Montreal Cognitive Assessment; PD-CRS, Parkinson's disease-Cognitive Rating Scale.

Acknowledgement

The authors are grateful to the two patients for their time and for sharing their subjective experience of hallucinations with us. We also thank Prof. Paul Krack, Prof. Ghika, Dr. Benoît Wicki, Dr. Fosco Bernasconi and Dr. Oliver A. Kannape for their review and critique of the manuscript.

Disclosures

Funding Sources and Conflict of Interest

The present observational report is part of a larger research project investigating hallucinations in Parkinson's disease. This project was supported by two generous donors

Appendix

advised by CARIGEST SA, the first one wishing to remain anonymous and second one being *Fondazione Teofilo Rossi di Montelera e di Premuda*; the Bertarelli Foundation; the Swiss National Science Foundation [grant number FNS 320030_188798], Parkinson Suisse Foundation; and Novartis Foundation [grant number 20A054]. The authors report no conflict of interest.

Financial Disclosures for the previous 12 months

The authors declare that there are no additional disclosures to report.

Ethical Compliance Statement

We confirm that we have read the Journal's position on issues involved in ethical publication and affirm that this work is consistent with those guidelines. The patients were recruited in Switzerland from La Tour Hospital (Geneva) and Geneva University Hospital. The two patients were identified in the setting of an ongoing study on hallucinations in PD. All patients gave written informed consent prior to participating in the study approved by the local ethics committee of the *Commission cantonale d'éthique de la recherche (CCER)* of Geneva (protocol n° 2019-02275).

References

1. Lenka A, George L, Arumugham SS, Hegde S, Reddy V, Kamble N, et al. Predictors of onset of psychosis in patients with Parkinson's disease: Who gets it early? *Parkinsonism Relat Disord*. 2017 Nov;44:91–4.
2. Ffytche DH, Creese B, Politis M, Chaudhuri KR, Weintraub D, Ballard C, et al. The psychosis spectrum in Parkinson disease. *Nat Rev Neurol*. 2017 Feb;13(2):81–95.
3. Fénelon G, Soulas T, De Langavant LC, Trinkler I, Bachoud-Lévi A-C. Feeling of presence in Parkinson's disease. *J Neurol Neurosurg Psychiatry*. 2011 Nov;82(11):1219–24.
4. Bernasconi F, Blondiaux E, Potheegadoo J, Stripeikyte G, Pagonabarraga J, Bejr-Kasem H, et al. Robot-induced hallucinations in Parkinson's disease depend on altered sensorimotor processing in fronto-temporal network. *Sci Transl Med [Internet]*. 2021 Apr 28 [cited 2021 Apr 28];13(591). Available from: <https://stm.sciencemag.org/content/13/591/eabc8362>
5. Takakusaki K. Functional Neuroanatomy for Posture and Gait Control. *J Mov Disord*. 2017 Jan;10(1):1–17.

Appendix

6.4 Neuroscience robotics for controlled induction and real-time assessment of hallucinations

Fosco Bernasconi ^{1,§}, Eva Blondiaux ^{1,§}, Giulio Rognini ¹, **Herberto Dhanis** ¹, Laurent Jenni ¹, Jevita Potheegadoo ¹, Masayuki Hara ², Olaf Blanke ^{1,3,*}

Affiliations

¹ Laboratory of Cognitive Neuroscience, Center for Neuroprosthetics & Brain Mind Institute, Faculty of Life Sciences, Swiss Federal Institute of Technology (EPFL), Geneva, Switzerland.

² Graduate School of Science and Engineering, Saitama University, Saitama, Japan.

³ Department of Clinical Neurosciences, Geneva University Hospital, Geneva, Switzerland.

[§] *Both authors contributed equally*

Corresponding Author

Prof. Olaf Blanke

Bertarelli Chair in Cognitive Neuroprosthetics
Center for Neuroprosthetics & Brain Mind Institute
School of Life Sciences
Ecole Polytechnique Fédérale de Lausanne (EPFL)
Campus Biotech, H4.3
Ch. des Mines 9
CH-1202 Geneva
E-mail: olaf.blanke@epfl.ch
Tel: +41 (0)21 693 69 21

Appendix

Abstract

Although hallucinations are important and frequent symptoms in major psychiatric and neurological diseases, little is known about their brain mechanisms. Hallucinations are unpredictable and private experiences, making their investigation, quantification, and assessment highly challenging. A major shortcoming in hallucination research is the absence of methods able to induce specific and short-lasting hallucinations, which resemble clinical hallucinations, can be elicited repeatedly, and vary across experimental conditions. By integrating clinical observations and recent advances in cognitive neuroscience with robotics, we have designed a novel device and sensorimotor method able to repeatedly induce a specific, clinically relevant hallucination: presence hallucination. Presence hallucinations are induced by applying specific conflicting (spatiotemporal) sensorimotor stimulation including an upper extremity and the torso of the participant. Another, MRI-compatible, robotic device using similar sensorimotor stimulation permitted the identification of the brain mechanisms of these hallucinations. Enabling the identification of behavioral and a frontotemporal neural biomarkers of hallucinations, under fully controlled experimental conditions and in real-time, this method can be applied in healthy participants as well as patients with schizophrenia, neurodegenerative disease, or other populations with hallucinations. The execution of these protocols requires intermediate-level skills in cognitive neuroscience and MRI processing, as well as minimal coding experience to control the robotic device. These protocols take ~3h to be completed.

Appendix

Introduction

Hallucinations are defined as aberrant perceptual experiences that take place in the absence of corresponding external stimuli. They can occur in a single (i.e. auditory) or multiple sensory modalities (i.e. auditory and visual) and their content may be simple or complex. Insight into the aberrant origin of hallucinatory experiences (i.e. recognizing the hallucination as a false perception by the patient) may be lost in some patients ^{1,2}. Hallucinations occur in 5-10% of the general population, in the absence of any diagnosed disease or disorder ³, are more frequent in the elderly ^{4,5}, and have been reported in up to 50% following bereavement ^{6,7}. Hallucinations are also of medical importance and are reported in various clinical populations. For example, hallucinations are prominent clinical features in patients with psychiatric and neurodegenerative diseases, such as schizophrenia ⁸, dementia with Lewy bodies (DLB) ⁹, and Parkinson's disease ¹⁰. However, the sensory modality, frequency, and severity of hallucinations differs across these conditions.

Schizophrenia is a psychiatric disorder with symptoms typically starting in late adolescence or early adulthood, and include psychotic symptoms such as hallucinations, delusions, and thought disorder, often requiring lifelong treatment ⁸. Hallucinations are among the most frequent symptoms in schizophrenia, and predominantly consist of auditory hallucinations (voices - defined as auditory verbal hallucinations (AVH); 60-80% of patients ¹¹), but other hallucinations such as visual hallucinations may also occur (simple visual patterns, people, animals; 30-50% ¹¹). A frequent, but less investigated hallucination in schizophrenia is the presence hallucination (PH; 46% ¹²; defined as the vivid sensation that another person is nearby when no one is actually present and can neither be seen nor heard ¹³).

In neurodegenerative disease, hallucinations are most frequently observed in DLB, the second most common neurodegenerative dementia following Alzheimer's disease ¹⁴⁻¹⁶. Visual hallucinations occur in about 80% of such patients and have been identified as one of the three core features for the diagnosis of DLB (also including parkinsonian motor symptoms and cognitive decline) ¹⁴. PH and AVH are less frequent and both are described in 20-40% of patients with DLB ¹⁷⁻¹⁹.

Although Parkinson's disease is primarily defined as a movement disorder (i.e. rigidity, tremor, bradykinesia), approximately half of the patients with Parkinson's disease suffer from hallucinations, which increase in frequency and severity as the disease progresses. Visual hallucinations, PH, and auditory hallucinations have been described in 30-60% ^{10,20}, 25-42% ^{21,22}, and 5-15% ¹⁸ of Parkinson's disease patients, respectively. Hallucinations in Parkinson's disease are associated with and have been argued to be putative predictors of major negative clinical outcomes such as delusions, cognitive decline and dementia, as well as earlier home placement and a higher risk of mortality. Moreover, there is growing consensus that hallucinations in Parkinson's disease form a continuum progressing throughout the disease ^{10,22,23}, with the typical scenario starting with the so-called minor hallucinations (PH, passage hallucinations (brief images of a person or animal passing sideways, within the peripheral visual field), visual illusions) with preserved cognition that evolves towards visual hallucinations and cognitive decline and dementia. Because of the earlier onset of minor hallucinations (i.e. they may occur in some cases even before motor symptoms ²⁴) and their likely transition towards visual hallucinations, it has been argued that minor hallucinations, including PH, may be an early indicator of a more severe form of Parkinson's disease ^{22,25}.

Appendix

Despite the elevated frequency and clinical importance of hallucinations in several neurological and psychiatric diseases (e.g. ^{10,11}), hallucination research is hampered by the low number of devices and experimental approaches for hallucination induction. Previous attempts to elicit hallucinations, such as drug-induced hallucinations or those induced by automatized visual stimulation (i.e. Flicker-induced hallucinations) failed to induce clinically relevant hallucinations (i.e. resembling those experienced by patients in daily life). These methods also do not induce specific hallucinations with clear onset and offsets, but rather complex multisensory hallucinations with highly variable onsets-offsets. Moreover, the intensity of such induced hallucinations may vary and can be associated with more severe alterations of consciousness (i.e. clouding of consciousness, delirium, agitation) that limit their impact for clinical, legal, and ethical reasons (for more details see section Comparison with other methods).

Hallucinations are subjective and private experiences and current methods to investigate and evaluate hallucinations are based on self-reports and interpretations of those reports by investigators, known to be associated with several limitations. Verbal reports and graphical depictions often only capture some aspects of conscious experience and are prone to the well-known participant and experimenter biases ^{26,27}, which may be exacerbated in patient populations. Reporting hallucinations is further affected by the fear of stigmatization, which often causes patients to refrain from reporting them openly to caregivers and clinicians ²³. The unpredictable and spontaneous appearance of hallucinations, often outside the hospital or research setting, further complicates their study. It is very unusual for a research to be able to investigate the hallucination of a given participant in real-time, that is while the person is undergoing a hallucinatory experience.

Another major shortcoming is that hallucination research is biased towards research in patient populations, despite the demonstration that they can also occur in healthy individuals ³. Investigating how hallucinations arise in the nervous system of healthy individuals, in parallel with patient studies, will improve our possibilities of understanding the actual mechanisms underlying hallucinatory perceptions; because healthy individuals do not have confounding medical variables present in most patients, such as medications, comorbidities, neuropsychological deficits, or impairments of consciousness ²⁸. Therefore, laboratory-based methods and procedures allowing the repeated, safe, controlled, and real-time induction of well-defined hallucinations are critical to determine their behavioral and brain mechanisms and to develop new diagnostic and prognostic tools for different disorders characterized by hallucinations.

Development of the protocol

By integrating clinical observations and recent advances in cognitive neuroscience in healthy participants with robotics, we have designed a new method able to induce a specific, clinically relevant hallucination in healthy participants: PH. Such robot-induced PH (riPH) were induced repeatedly and in real-time in the laboratory or in different clinic settings and were controlled by a robotic device across several experimental conditions. The method for riPH was developed and first investigated in healthy participants ^{29,30,31} and subsequently tested in patients with Parkinson's disease ³¹ and schizophrenia ³².

Previous work suggested that hallucinations such as AVH, but also related sensations such as loss of agency and passivity experiences, may be associated with altered sensory prediction of self-generated actions ^{33,34} and the inability to update or increase the precision of this prior ². Impairments in sensory

Appendix

prediction and sensorimotor integration have also been associated with PH in neurological patients^{29,35,36}. Furthermore, invasive electrical stimulation (in a patient undergoing presurgical epilepsy evaluation) of the temporoparietal cortex, which is associated with sensorimotor integration, resulted in the repeated and the controlled induction of PH³⁷, and damage to temporoparietal cortex and related regions has been associated with PH and sensorimotor deficits²⁹, corroborating the important role of altered sensorimotor processing in PH of neurological origin.

By integrating these clinical observations with the human neuroscience of own body perceptions in healthy participants (i.e.^{38–40}) and advances in robotics, we developed the following robotic method and sensorimotor procedure for PH induction (Figure 1). Participants are asked to perform repetitive movements to operate a robot placed in front of them (front-robot), which is combined with a back-robot providing tactile feedback (i.e. touches or strokes) to the participants' backs⁴¹. When such sensorimotor stimulation and tactile feedback are delivered with a delay between a participant's hand movements and the tactile feedback they receive on their back (i.e. asynchronous condition), they reported riPH (Figure 2A-B)^{29,31,32,42}. We further observed that an increase in such asynchronous delays (increase in sensorimotor conflict), increases the probability to induce a riPH³¹ (Figure 2C). However, when the tactile feedback is delivered synchronously with the participant's hand movements (synchronous control condition; but with a spatial conflict between the movement in front and touch feedback delivered on the back) participants did not experience riPH, but instead reported control sensations (e.g. illusory self-touch), compatible with previous works on own body perceptions^{38,39,43–46}.

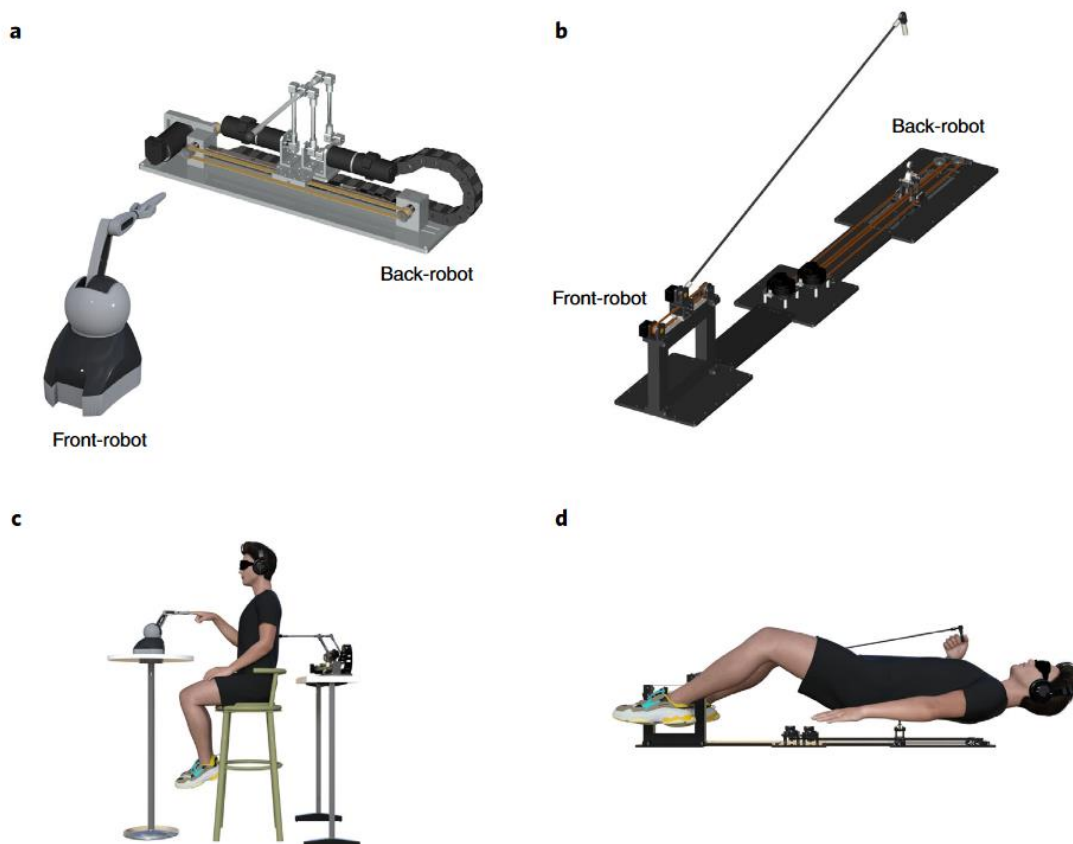


Fig. 1 | Robotic setup. **a**, Behavioral robotic device. **b**, MR-compatible robotic device. **c**, Behavioral setup allowing the induction of riPH, depending on the sensorimotor condition. **d**, MR setup allowing the induction of riPH, depending on the sensorimotor condition. The participant is lying on the back-robot (the plank on which the participant is laying on has been removed from the illustration to avoid hiding the back-robot).

Appendix

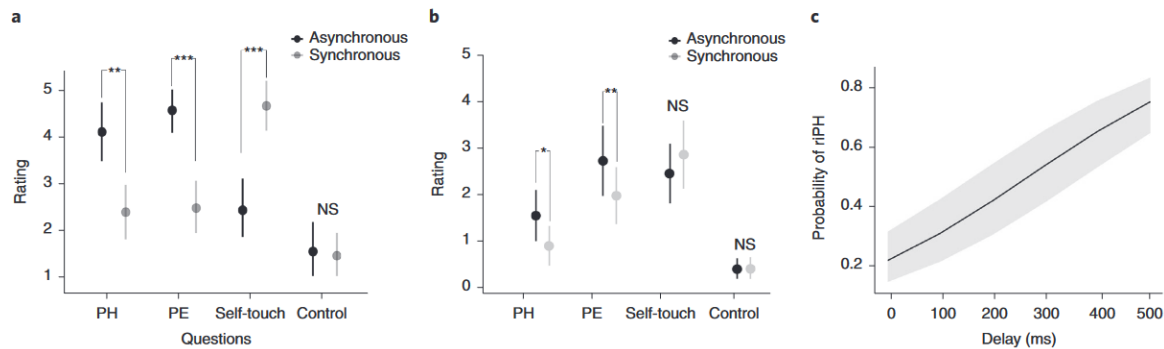


Fig. 2 | Behavioral results of the robotic sensorimotor stimulation. **a**, Questionnaires' results, showing a significant induction of rIPH (and other illusions) during the condition with strongest sensorimotor conflict (results from ref. ²⁹). The dots indicate the mean ratings; the error bars indicate the 95% confidence interval. Data were analyzed fitting cumulative link mixed model. **b**, Questionnaire results using the MR-compatible robotic device (results from ref. ³¹). The dots indicate the mean ratings; the error bars indicate the 95% confidence interval. Data were analyzed fitting cumulative link mixed model. **c**, rIPH is modulated as a function of the sensorimotor delay. Data were analyzed fitting a binomial generalized linear mixed models. The thicker line indicates the mean of the fitted models; the shaded area indicates the 95% confidence interval. The questions are the following: (1) PH: 'I had the impression that someone was standing close to me, behind or next to me'; (2) passivity experience (PE): 'I had the impression that someone else was touching my back'; (3) self-touch: 'I had the impression that I was touching my back'. Asterisks indicate significant differences across sensorimotor conditions, for the indicated question. NS, not significant.

More recently, the robotic method and procedure were adapted for clinical research and it was shown that patients with Parkinson's disease and preexisting symptomatic PH (PD-PH) have a higher sensitivity to the rIPH method, compared to patients with Parkinson's disease without pre-existing symptomatic PH. In addition, about 40% of the PD-PH patients experienced rIPH to one side of their body and not on their back (i.e. where the robotic tactile feedback was actually applied). Some patients directly likened the rIPH with their symptomatic PH. These findings reveal relevant phenomenological similarities between such experimentally controlled rIPH and the same patients' preexisting symptomatic PH that are often experienced by patients to one side of their body and of the same intensity: The conscious experience associated with rIPH shares characteristics with spontaneously occurring symptomatic PH ³¹.

Overall, these results demonstrate that specific sensorimotor conflicts, including bodily signals from the arm and trunk, are sufficient to induce mild to moderate rIPH, linking PH to the misperception of the source and identity of sensorimotor signals of one's own body. These data also show that rIPH share phenomenological characteristics with symptomatic PH. Moreover, rIPH were induced safely and repeatedly and were robotically controlled across experimental conditions, which has not been achieved before. This enables the repeated real-time assessment of a specific hallucination compared to standard assessments based on the retrospective evaluation of spontaneously occurring hallucinations that may have occurred days (or months before) and only a limited number of times. The assessment of rIPH can be based on established procedures in human neuroscience, including among others hallucination proneness (measured with questionnaires or implicit measures (see below for more details)), qualitative measures (phenomenological description of the hallucination - e.g. spatial location), and behavioral readouts (e.g. drift in self-location, numerosity estimation).

By adapting the robotic technology ²⁹, sensorimotor method, and experimental protocol to the MRI environment it has also been possible to identify the brain network associated with rIPH in greater detail and in healthy participants ³¹. Robotic sensorimotor stimulation activated a network of brain regions consisting of the inferior frontal gyrus (IFG), anterior insula, medial prefrontal cortex, and the posterior part of the superior temporal sulcus (pSTS), especially in the asynchronous condition (Figure 3). Moreover, additional research showed that these regions were partially overlapping with brain

Appendix

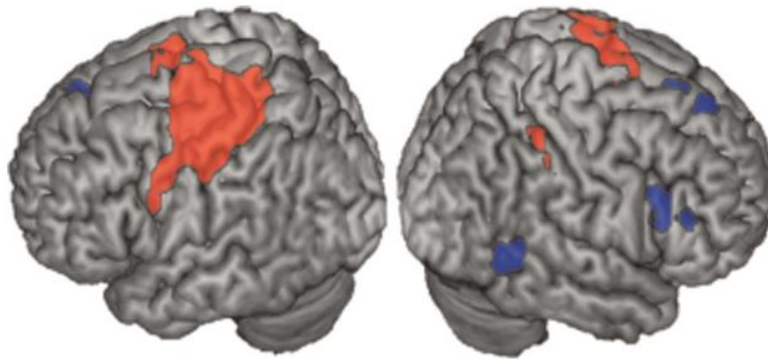


Fig. 3 | Brain network of riPH. Brain activations in the asynchronous > synchronous contrast (blue), showing the brain regions more activated in the asynchronous condition compared with the synchronous condition (results from ref. ³¹). In red, the conjunction analysis asynchronous > [motor + touch] \cap synchronous > [motor + touch] is shown (results from ref. ³¹). This conjunction contrast highlights the brain regions involved in the spatial sensorimotor conflict between the sensorimotor movement of the hand in front space and the feedback in the back, independent of the sensorimotor delay.

regions associated with the symptomatic PH of neurological origin (caused by stroke or focal epilepsy) ²⁹. This anatomical overlap (between the riPH fMRI network in healthy participants and the PH network in neurological patients suffering from non-parkinsonian neurological illnesses) allowed us to define a common cortical PH-network that included the pSTS, the inferior frontal gyrus (IFG), and the ventral premotor cortex (vPMC). We confirmed the relevance of this latter network in patients with Parkinson's disease by observing a disruption within this common PH-network, describing altered connectivity between the left IFG and left pSTS. This disruption was only found in Parkinson's disease patients with preexisting symptomatic PH, but not in Parkinson's disease patients without such PH, revealing pathological cortical sensorimotor processing related to a specific hallucination in Parkinson's disease.

Collectively, these results on riPH show that the robotic procedure can be used to induce specific hallucinatory states in real-time, repeatedly, and under controlled conditions in patients with Parkinson's disease and in healthy individuals. The hallucinatory states are associated with a specific brain network which is disrupted in Parkinson's disease patients with preexisting symptomatic PH. The possibility to study a clinically relevant hallucination in healthy individuals will further allow research to progress faster (avoiding limitations due to patient recruitment and testing) and to identify the mechanisms and brain networks without clinical confounds (medications, comorbidities, neuropsychological deficits, or impairments of consciousness) present in many patient populations. Investigating a specific hallucination in real-time also allows reducing confounds arising from recollecting and interpreting complex subjective reports about hallucinations that have occurred outside the controlled laboratory environment and also facilitates the real-time study of the neural correlates of specific hallucinations and their potential as hallucination biomarkers, which is very challenging for spontaneously occurring hallucinations (which are often associated with other clinical confounds).

Applications of the method

Schizophrenia, a psychiatric disorder with a worldwide prevalence of approximately 1% ⁴⁷, is characterized by psychotic symptoms require lifelong treatment ⁸. AVH are the most prevalent

Appendix

hallucination experienced in schizophrenia and have been suggested to arise from abnormal sensorimotor prediction^{2,48} and, among others, the inability to update the priors associated with these sensorimotor signals².

We recently adapted our robotic protocol that induces riPH to the investigation of the relation between sensorimotor stimulation and auditory voice perception⁴⁹. Evidence from healthy participants⁴⁹ showed that conflicting sensorimotor stimulation applied with the robotic method resulted in systematic and predictable errors in voice perception. Similarly, results from psychiatric patients with early psychosis showed a selectively reduced accuracy in auditory-verbal (self-other) discrimination during conflicting sensorimotor stimulation applied with the robotic method, especially in individuals suffering from so-called first-rank symptoms, including AVH and delusions³².

These results confirm an association between altered sensorimotor processes and auditory-verbal misattribution that is typically observed in schizophrenia. Recent neuroimaging results corroborated and extended these behavioral findings by showing that the common PH-network (as described above) in a large group of patients with early psychosis was characterized by alterations of the common PH-network and its connectivity with Heschl's gyrus, a brain region linked to auditory perception and auditory-verbal hallucinations⁵⁰. More work is needed to directly investigate the behavioral and brain mechanisms of symptomatic PH and riPH in schizophrenia¹² and how the sensorimotor PH mechanisms relate to those of AVH, delusions, and related cognitive functions in schizophrenia, as recently shown for the psychosis-related phenomenon of thought insertion⁴².

DLB, a neurological disorder, and the second most common dementia subtype⁵¹, is typically characterized by hallucinations, which are one of the three defining symptoms for DLB diagnosis. Despite the high prevalence of PH in DLB very little is known about the brain mechanisms underlying this specific hallucination in DLB, as past research has focused on visual hallucinations. By adapting our robotic sensorimotor procedure to patients with DLB we may be able to further elucidate the (aberrant) brain mechanisms of PH and also determine overlapping and distinct PH mechanisms by comparing PH in patients with DLB versus those of patients with Parkinson's disease and healthy controls³¹. This would provide us with additional information on the neural causes of symptomatic PH, how this depends on disease type, and eventually aiding differential diagnosis among PH of neurodegenerative origin. Recent work allowed us to demonstrate that patients with DLB and symptomatic PH are characterized by alterations in the common PH network⁵². Additional research is required to investigate whether these PH-related alterations precede those associated with visual hallucinations (one of the key symptoms of DLB), as observed for Parkinson's disease. Finally, hallucinations, including PH, in patients with Parkinson's disease have been associated with cognitive decline¹⁰ and riPH will help the investigation and such associations in patients with DLB, requiring longitudinal clinical research.

Hallucinations are not only associated with psychiatric or neurodegenerative disease but are also experienced by the general population in the absence of any diagnosed disease or disorder³. For instance, 20% of individuals over 60 years old have reported experiencing hallucinations^{4,53}, and PH have been observed in 5% of those individuals. Whether these events are indicative of the onset of mild cognitive decline or are limited to specific events such as traumatic events (bereavement⁵³) is still unknown. However, tremendous research current efforts are dedicated to the early detection of dementia, its distinction from normal age-related decline, and the early distinction between cognitive decline in Alzheimer's disease versus DLB and other dementias^{54,55}. We suggest that the adaptation

Appendix

and further integration of the described riPH procedure with other psychophysical, neuropsychological, clinical, and neuroimaging techniques may help early detection and differentiation of clinical early dementia from normal aging, by adding a hallucination-psychosis dimension to current cognitive evaluations.

Comparison with other methods

Alternative methods and procedures able to induce hallucinations by using experimental methods have been described, but are sparse and suffer from several limitations. Flicker-induced hallucinations^{56,57}, Ganzfeld-induced hallucinations⁵⁸, or visual/sensory deprivation^{59,60} are among some of the most well-known methods. Although these methods are relatively effective in inducing hallucinations in healthy participants, they have been mostly used to investigate visual hallucinations and have the following limitations. The experimental control over the specific content of such induced hallucinations, as well as their onset and offset are very low. Thus, there are multiple characteristics of the induced visual hallucinations (color, form and motion) that are very heterogeneous, which in addition typically change unpredictably over time. Such methods generally do not allow for the repeated induction of a hallucination for several times within a short time period. Moreover, these methods do not seem to induce hallucinations that are comparable with those experienced by neurological or psychiatric patients (e.g. for flickering stimuli see⁶¹). Other procedures have employed conditioning-induced hallucinations by pairing visual and auditory stimuli to evoke hallucinations in one or the other modality^{62,63}. While this method is more reliable than the other procedures presented in this section, the time needed to induce such hallucinations is relatively long, the hallucinations induced are restricted to simple stimuli (e.g. tone or flash), and rarely match the phenomenology of clinical hallucinations.

Other methods to induce and study hallucinations in healthy participants exist, but many of these have potentially serious side effects. For instance, it is possible to induce a range of complex hallucinations by pharmacological interventions (for a review see⁶⁴). Substances like ketamine, lysergic acid diethylamide (LSD), phosphoryloxy-N,N-dimethyltryptamine (psilocybin), or N,N-dimethyltryptamine (DMT) have been investigated intensively in the past and have received renewed interest^{64,65}. By combining such pharmacological interventions with behavioral, electroencephalography and neuroimaging data it has been possible to identify some of the brain mechanisms associated with drug-induced hallucinations. For instance, visual hallucinations have been associated with abnormal activation of the 5-hydroxytryptamine (serotonin) type 2A (5-HT_{2A}) receptors⁶³. These findings are of high relevance for the development of new pharmacological treatments aiming at treating hallucinations. However, pharmacological studies suffer from similar shortcomings as noted above for flicker- and Ganzfeld-induced hallucinations (i.e. onset, offset, repetition in a short time frame, content, etc.). Moreover, higher doses may induce very intense hallucinations that may be associated with impairments of consciousness, making the real-time description and investigation (very) challenging. Lower doses, on the other hand, likely fail to induce hallucinations⁶³. In addition, in most cases of drug-induced hallucinations, a large range of visual, auditory, somatosensory and other hallucinations are generally reported simultaneously, making the investigation of specific drug-induced hallucinations difficult. Drug-induced hallucinations also raise several ethical concerns (such as inducing uncontrolled psychotic episodes in healthy participants,

Appendix

providing psychoactive drugs to patients without being able to predict the reaction to the substance, etc.) that are not, or at least less, present in other hallucination-inducing techniques.

Finally, another hallucination method is based on invasive electrical stimulation of the brain, but can only be applied in highly specialized clinical settings and proposed to specific patient populations (for example to epilepsy patients undergoing invasive presurgical evaluation). Thus, electrical stimulation of specific brain regions has been shown to induce PH³⁷, simple and complex hallucination⁶⁶, visual hallucinations⁶⁷, AVH⁶⁷, vestibular hallucinations^{68,69}, and somatosensory hallucinations⁷⁰. Hallucinations induced by invasive brain stimulation can be considered the gold standard of linking brain circuits to hallucinations, due to their high spatial resolution, full control of the stimulation and hallucination onset and offset, and the repeated induction of highly specific hallucinations, including PH³⁷. However, this powerful approach is limited to rare patient explorations, requires a neurosurgical intervention, and is not applicable for systematic experimental research in healthy people and other non-neurosurgical populations.

Experimental design

To experimentally investigate riPH we have developed several distinct experimental procedures i) based on comparing asynchronous versus synchronous sensorimotor stimulation conditions, ii) additionally manipulating somatosensory force feedback (“virtual back” in front of the participant), and iii) based on the determination of sensorimotor delay dependency.

General experimental procedure (for behavioral and brain imaging studies)

Robotic sensorimotor stimulation is administered with a robotic system^{30,41} consisting of two robotic components (front-robot, back-robot) that we previously have used to induce PH^{29,31,32,42}. Participants are asked to insert their index finger in the front-robot and carry out repeated poking movements while they receive tactile cues on their back, delivered by the back-robot (Figure 1). Such sensorimotor stimulation includes motor, tactile, and proprioceptive signals from the upper limb moving the front-robot and additional tactile signals from the back-robot. At the beginning of the experiment, after having explained and inspected the robotic setup, participants are instructed to perform several poking movements (in the synchronous condition) to become familiar with the robotic device, the movements, and the tactile feedback. During the experimental conditions, participants are instructed to perform arm movements at a frequency of approximately 1Hz (1 poke per second), and to explore their whole back with their movements (Figure 4). Participants must be instructed not to count the time between pokes nor the number of pokes, but do the movement as learned during the training. During the sensorimotor stimulation, participants are either blindfolded or asked to keep their eyes closed for the remainder of the experiment. In addition, participants are exposed to continuous pink noise through headphones.

Appendix



Fig. 4 | Example of back exploration done for the behavioral experiment. Each dot indicates a point in which the back-robot applied touch on the back of the participant.

Spatio-temporal sensorimotor stimulation

This experimental protocol (Figure 5) allows to investigate riPH as well as other bodily experiences (self-touch, passivity experience, etc.)^{29,31,32,42}. This protocol (Spatio-temporal sensorimotor stimulation) has the advantage to be of very short duration and can therefore be done by most participants, including patients where time may be limited. The back stroking is applied either synchronously (0 ms delay; control condition) or asynchronously (500 ms delay) with respect to the movement performed by the participant with the front-robot.

The versatility of the robotic system provides the possibility to apply the touch feedback from the back-robot not only to the back, but also to other body parts (e.g. the hand). Because PH is associated with a misinterpretation of the global (trunk centered) self²⁹, stimulating the hand should not induce riPH, but it can induce other illusions such as loss of agency or passivity experience (i.e. the illusion of being touched by someone else)^{48,71}. It can therefore be used as control experimental conditions. The delays of 0ms and 500ms have been motivated by prominent work in sensorimotor and multisensory integration and sense of agency, as well as the forward model of motor control that have been applied to bodily illusions^{43,72–76} and hallucinations⁴⁸.

For behavioral experiments, the sensorimotor stimulation lasts 2 consecutive minutes for each of the experimental conditions (synchronous and asynchronous). In total, the experiment generates two datasets (1: Synchronous – 2 minutes; 2: Asynchronous – 2 minutes), for a total of 4 minutes, with approximately 120 pokes per dataset. For fMRI experiments, the sensorimotor stimulation is reduced to 30 seconds. The order of the conditions must be randomized across participants. The assessment

Appendix

of riPH and other specific bodily illusions is done through explicit measures (questionnaires see Supplementary Table S1 and ^{29,31,32,42}) or implicit measures such as self-location ²⁹ or numerosity estimation ²⁹ that are administered immediately after each condition. Self-location is quantified using the mental ball-throwing task, during which participants are asked to estimate (by pressing a button) the time that a ball they holding in their hands would take to reach the wall if they were to throw it. Numerosity estimation is quantified asking participants to judge how many people do they feel close. Importantly, the participant must stay blindfolded during these tasks.

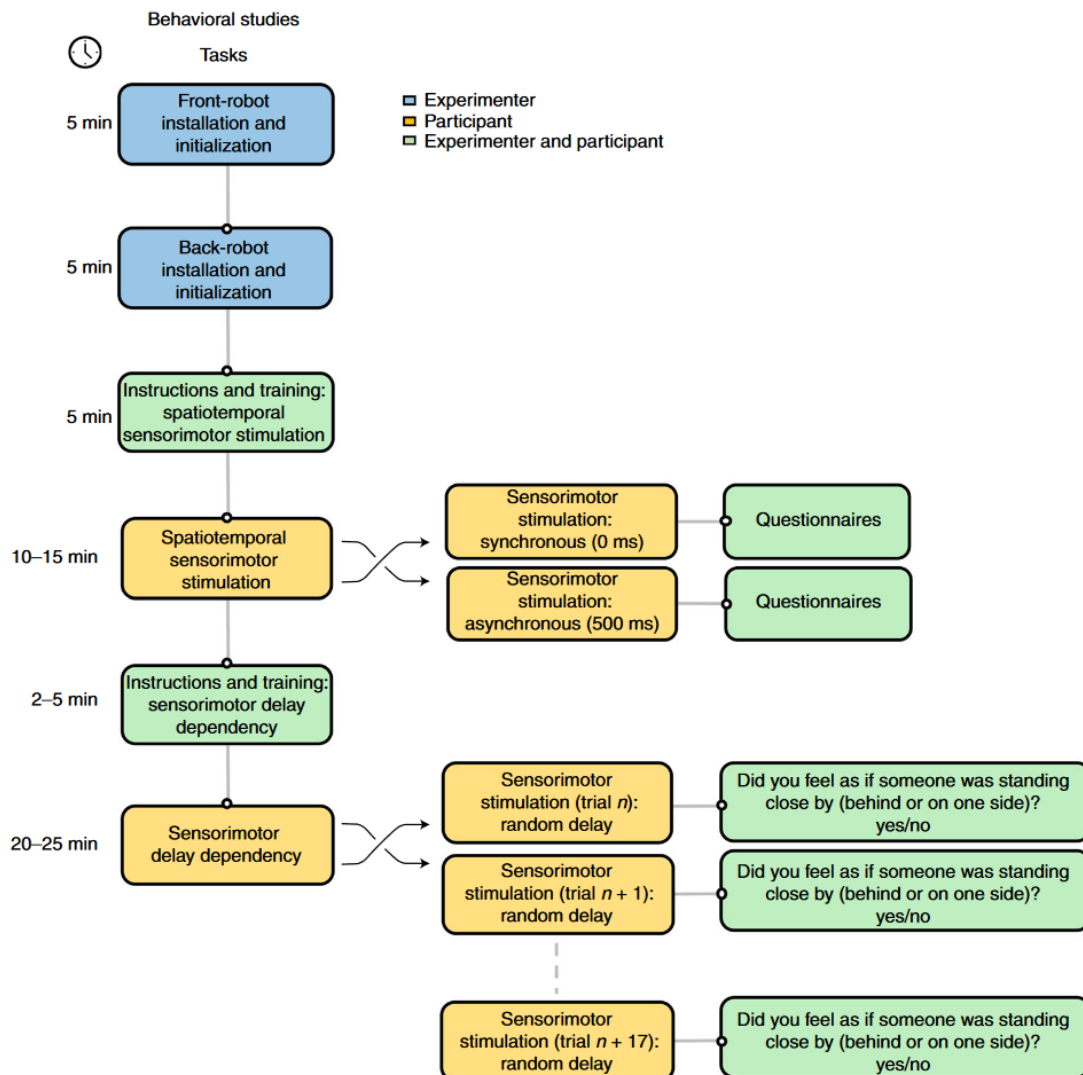


Fig. 5 | The experimental workflow for behavioral experiments is divided into three principal stages. First, initialization of the front and back-robot and installation of the participant is performed by the researcher (blue). Next, participants (yellow) will be given instructions and a short training session to familiarize themselves with the robotic setup and its use. Before starting the first experiment (i.e., ‘Spatiotemporal sensorimotor stimulation’), participants are asked to wear headphones and a blindfold. In our previous work³¹, the synchronous (0 ms) and asynchronous (500 ms) conditions were each performed for 2 min, each followed by a short questionnaire to assess the intensity of PH and other mental states. The arrow crossing indicates that the two experimental conditions must be randomized in order across participants. Finally, the researcher will instruct the participant on how to perform the second part of the study (i.e., ‘Sensorimotor delay dependency’). In our previous work³¹, the ‘Sensorimotor delay dependency’ task was repeated three times (18 trials per block, delays 0–500 ms). In total, each delay (between 0 and 500 ms, steps of 100 ms, six delays tested) was randomly repeated nine times. Thus, the total number of trials across all sessions and delays was 54. Experimental parameters can be adapted on the basis of the research questions and needs.

Appendix

Somatosensory force feedback

This experimental protocol extends the Spatio-temporal sensorimotor stimulation protocol and allows testing how the induced PH depends on the the force that participants apply at the front-robot and consequently receive on their back by the back-robot. The front-robot allows manipulating the somatosensory force feedback (force conditions) that is applied on the back of the participant. Specifically, using an impedance control, a “virtual wall” with a stiffness of 1.0 N/mm can be created with the front-robot. The “virtual wall” consist in a force applied by the front-robot in the opposite direction to the force done by the arm extension, this occurs only when the front-robot reaches the predefined coordinate in which the “virtual wall” is located. Hence, the participant will be able to extend the movement to a predefined distance (i.e. position of the “virtual wall”, after which the robot will stop the movement as if there was a “virtual” wall beyond which no movement is possible. The force feedback also includes a mechanical stop occurring (synchronously or asynchronously) to the touch that the participant received on the back ²⁹. As for the “Spatio-temporal sensorimotor stimulation” riPH can be assessed with implicit and explicit measures.

Sensorimotor delay dependency

This experimental protocol (Figure 5) allows the assessment of the proneness to induced hallucinations (e.g. PH), and how this depends on the degree of the delay of the applied sensorimotor conflict. Compared to the Spatio-temporal sensorimotor stimulation experimental protocol described above, here we use a two-alternative (Yes/No) experimental paradigm, in which participants are asked to perform the poking movement (and receive the stimulation on the back) and then report (verbally), on a trial-by-trial basis, whether they experienced the hallucination or not. [question: “Did you feel as if someone was standing close by (behind you or on one side)?”], without opening their eyes.. Compared to the experiment mentioned above, the trials here are shorter, lasting only about 10 seconds. Accordingly, here we can repeat the hallucination induction over many more trials as well as test several controlled experimental parameters (such as the delay, the position of the tactile feedback on the back, the presence/absence of the virtual wall, the duration of the trials, etc.). In our previous work testing the effect of delay ³¹, in total, participants were exposed to 3 experimental blocks and each block was composed of 18 trials consisting in random order of the different delay conditions. In total, each delay (between 0-500ms, steps of 100ms, 6 delays tested) was randomly repeated 9 times. Thus, the total number of trials across all sessions and delays was 54). We note that the number of trials (as well as the other parameters) can be adapted based on the experimental question and/or the population tested.

On each sensorimotor stimulation trial, the temporal delay between the movement and the stroking on the back is randomly chosen between 0 and 500 ms (with 100 ms steps) ³¹. Larger delays (any delay) are also possible (maximum 3998ms). The beginning of each trial in the delay protocol of our method is indicated by an acoustic signal (e.g. 400 Hz tone, 100 ms duration) and after this tone, the participant will start performing the poking movements. Once the number of pokes reaches the desired number (automatically counted; for example six ³¹), two consecutive tones (400 Hz, 100 ms duration) indicate to the participant to stop moving the robot and to verbally answer Yes/No to the riPH question. To limit possible confounds, the investigators must always be placed as far as possible from the participant, and in front (not on the back or side) during the whole experiment. The order of the delays must be fully randomized. All the experimental parameters (e.g. delays, number of repetitions, etc.)

Appendix

can be adapted to the experimental needs. Similar to what is described above, we can manipulate the somatosensory force feedback and the body part stimulated.

Using different delays (presented randomly) has several advantages and provides the possibility to integrate several features from psychophysics into hallucination research and identify:

- i) the type of association between delays and the induced presence hallucination,
- ii) a difference in sensitivity to the sensorimotor task between groups of participants (e.g. groups of patients) as indicated by a difference in slope of the fitted regression,
- iii) a difference in the propensity to reporting “yes” (bias) would be reflected as a main effect of group (rather than a difference in slope).

fMRI protocol

To experimentally assess the brain mechanisms of PH (Figure 6), we use an MR-compatible robotic system³⁰ to generate sensorimotor conflicts in a similar manner as described above, but adapted to the supine body position in the scanner. The front-robot consists of a rod attached on a revolution joint, itself attached to a slider allowing the participant to move in the X (along the body) and Z (towards the body) directions. The movement of the carbon-fiber rod is electronically translated into movements of the back-robot. The back-robot is composed of an effector that touches the participant on its back through stroking and tapping movements. Furthermore, its shape was adapted to the space inside the scanner and a wooden mattress structure was specifically designed with a central slit to allow the tip of the back-robot to touch the back of the participants. The performance of the robotic system has been previously validated inside a 3T and 7T MR scanner with a vendor-supplied spherical phantom (Siemens)³¹.

Comparable to the behavioral experiments described above, participants are blindfolded and are wearing headphones in order to receive the auditory cues to start (one tone) and stop the movement (two consecutive tones). During the fMRI session, in addition to the experimental conditions (see below for details) the anatomical T1 weighted image is also acquired for each participant (Figure 6).

Spatio-temporal sensorimotor stimulation

Participants are asked to move the front-robot in a self-paced manner while receiving sensory feedback on their back either synchronously to the movement (synchronous condition) or with a delay of 500 ms (asynchronous condition). Depending on the research question, different delays can be used in the paradigm. The fMRI experimental design is a task-based block design. Participants perform two runs of 12 min each, during which they repeatedly have to move the front-robot for 30 s (16 repetitions per condition in total), followed by 18 s of rest. Conditions are presented in a randomized order.

Appendix

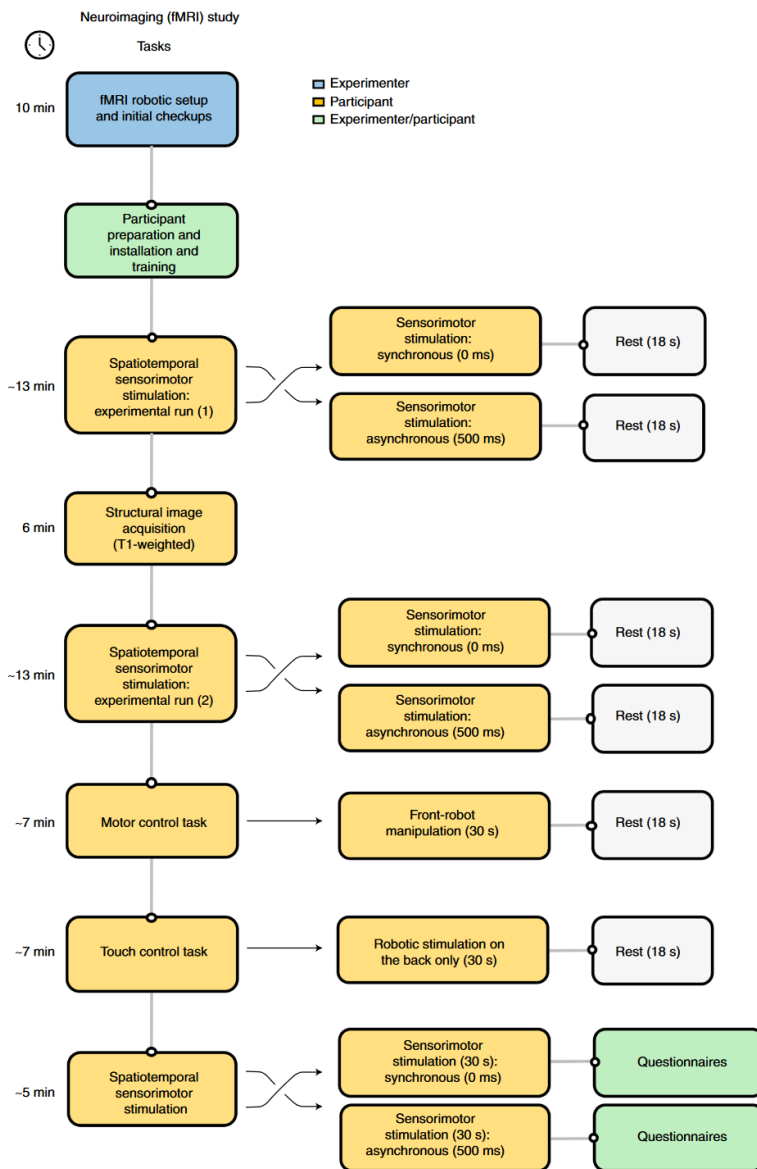


Fig. 6 | Experimental workflow for the fMRI experiment.

First, the fMRI robotic system is installed and initialized by the researcher (blue). Then participants are laid (with earplugs, headphones and blindfold) in the MR bed on a mattress below which the back-robot is installed. Next, participants (yellow) receive the instructions and short training with the robotic setup, before starting the first experimental run (i.e., 'Spatiotemporal sensorimotor stimulation'). In our previous work³¹, the synchronous (0 ms) and asynchronous (500 ms) conditions were given in randomized order and lasted 30 s each, followed by rest. The arrow crossing indicates that the two experimental conditions must be randomized in order across participants. Each condition was repeated eight times. The anatomical image is then acquired followed by a second experimental run (i.e., 'Spatiotemporal sensorimotor stimulation'). The control tasks (i.e., in the motor control task, participants are only moving the front-robot without receiving any sensory feedback on the back; in the touch control task, participants do not perform any movement but are touched on the back) are acquired (eight repetitions for each task). Finally, after the fMRI acquisition is terminated (but participants are still in the scanner), the participant will be asked to perform again 30 s of movement in each condition (i.e., asynchronous and synchronous) and will be asked to answer verbally to a short questionnaire that assesses the intensity of PHs and other mental states. Experimental parameters can be adapted on the basis of the research questions and needs.

Two additional control conditions are recorded:

- i) in a touch control task only the sensory feedback is applied (touch only; no movement done by the participant),
- ii) in a motor control task only the movement is performed (no touch is applied on the back).

In the motor control task, participants are asked to manipulate the front-robot in the same manner as during the two experimental runs (with the asynchronous and synchronous condition), however, in this condition they do not receive any tactile feedback on the back. Conversely, during the touch control task participants do not perform any movement with their right hand but receive tactile feedback on their back (from a pre-recorded sequence).

The experiment always starts with the two experimental runs and continues with the touch control tasks and ends with the motor control task. For each of the robotic stimulation periods, the movement parameters of the robotic system along the X and Z axis are recorded at a frequency of 100 Hz.

Appendix

From this, there would be a minimum of four datasets of functional images (two experimental runs (in the example used consisting of 320 functional images each, and two control tasks consisting of 160 functional images each and one t1 weighted images resulting from each experiment.

Expertise needed to implement the protocol

A mechatronic engineer (engineer trained in integrating mechanical and electronic engineering systems, which includes robotics and computer systems) with some experience in cognitive neuroscience might be necessary for the building and maintaining of the robotic devices. Details on how to build the robotic devices have been previously published (for the behavioral robot see ⁴¹, for the MRI robot see ³⁰).

Materials

Participants

Any individual able to perform the task can be enrolled in the studies.

CRITICAL Obtain informed consent from the participants (in agreement with local regulations) prior to starting any part of the experimental procedure. For the MRI protocol, participants must not present any contraindications (see Supplementary Table S2).

CRITICAL Obtain informed consent from the participants and conduct an anamnestic interview to verify the inclusion criteria (those can vary depending on the researchers interest) and collect the demographic data (age, sex, handedness through the Edinburg Handedness Inventory (Oldfield, 1971)).

Equipment

Behavioral protocol

A schematic illustration of the setup is shown in Figure 1A.

- Front- and back-robot ⁴¹
- Desktop computer with Visual Studio 2012 (Microsoft) installed
- A custom-made controller (motor driver) for controlling the robotic device with serial communication function
- Headphones
- Blindfold
- Questionnaire (Supplementary Table 1)

MRI protocol equipment A schematic illustration of the setup is shown in Figure 1B and 1D.

Appendix

CRITICAL All neuroimaging data must be organized according to the BIDS specifications ⁷⁷.

- MRI scanner, the protocol has been implemented using a 3 Tesla Siemens Magnetom Prisma MRI scanner using a whole-brain echo-planar imaging (EPI) sequences measuring the BOLD activity ³¹.
- Sequences: For the functional images, we used a whole brain echo-planar imaging (EPI) with the following parameters: 43 continuous slices, FOV=230mm, TR=2.5s, TE=30ms, flip angle=90°, in-plane resolution=2.5x2.5mm², slice thickness=2.5mm. We used a 64-channel head-coil. The two experimental runs contained 320 volumes and the control conditions (in which there is only 8 repetitions per condition) included 160 volumes. For each participant, an anatomical image was recorded using a T1-weighted MPRAGE sequence (TR=2.3s, TE=2.32 ms, Inversion time=900ms, flip angle=8°, 0.9mm isotropic voxels, 192 slices per slab and FOV=240mm).
- MR-compatible front- and back-robotic system ³⁰
- Desktop computer with Visual Studio (Microsoft; and platform toolset V110) installed and a data acquisition card (NI PCIe-6323, National Instruments) for controlling the robotic device
- The computer should also have Matlab 2020 with the Instrument Control toolbox (toolbox part of Matlab) and the psychtoolbox installed (freely available).
- Headphones
- Blindfold

Software

- For the preprocessing, the SPM12 toolbox (Wellcome Department of Cognitive Neurology, Institute of Neurology, UCL, London, UK) running in Matlab is needed (<https://www.fil.ion.ucl.ac.uk/spm/software/spm12/>)

CRITICAL This protocol is intended to be run on a modern workstation running an up-to-date Microsoft Windows

CRITICAL Scripts to control the robotic device are provided on Gitlab: <https://gitlab.epfl.ch/fbernasc/roboticsph>

Procedure 1 – Response measured by results of questionnaire.

CRITICAL The procedure does not require a specific directory structure or file-naming convention. However, users should modify the scripts to conform to their naming conventions and preprocessing pipelines.

Behavior

CRITICAL It is important that during the experimental part, the researcher(s) are in front of the participant (not on the back or the side, to avoid confounds on the illusions), at a minimal distance of

Appendix

2 meters. The back-robot should be placed against a wall, making it impossible that someone is located on the back. The distance between the left and right sides of the robot to the wall should be equivalent (to avoid confounds on the reported location of the riPH). Although the participant will wear a blindfold, if possible, the light should be dimmed. The chair must be without a backrest (or only with a very low backrest to not block the touch applied by the back-robot) and have a height that allows the participant to easily touch the whole back surface with the back robot. The initial position of the participant, the front- and back-robot are illustrated in Figure 1C.

Robotic setup and initial checkups [Timing: 10 minutes]

1. Install and initialize the front-robot (Touch USB (3D systems)).
 - Connect the USB cable to the computer (or Ethernet cable if you don't use the USB version of the PhantoM)
 - Connect the power supply
 - Open the Touch Setup application and follow the instructions displayed on screen ? Troubleshooting
 - Press the Save configuration button, the software close automatically
 - <Troubleshooting>
2. Install and initialize the back-robot.
 - Connect the USB cable, the power cable, and switch the electric box ON
 - Make sure you screw the correct rod at the back of the robot (depending on if you want to do a back or hand poking experiment) and open the corresponding controller application (HapticGroup_BackPocking_V5.1, or later version if available; can be found here: <https://gitlab.epfl.ch/fbernasc/roboticsph//Robots/Behavior/>) to calibrate the back-robot position according to the participant.
 - Make sure the front-robot and the back-robot are in their resting position (Supplementary Figure S1) and make sure the subject chair (and the subject) are not too close to the robot to avoid blocking its initialization motion
 - Start the initialize process ("Sampler (S)" then "ON" then in the pop-up window press "Initialization"). ? Troubleshooting
 - Press the "i" key. The robot starts moving to find the hard stops, and then it goes to its initial position. ? Troubleshooting
 - Once the robot stops the movement Stop the the initialize process ("Sampler (S)" then "OFF")
 - Start the initialize process ("Sampler (S)" then "ON" then in the pop-up window "Experiment")
 - Select "Condition 1" for heaving the virtual wall or "Condition 3" for no virtual wall (see Somatosensory force feedback)
 - If needed adjust the height of the robot (unscrew the vertical moving support and push the complete robot upward/downward then screw it back)

At the end of the whole experiment Stop the the initialize process ("Sampler (S)" then "OFF")

Appendix

Spatio-temporal sensorimotor stimulation

[Setup • Timing 5 minutes]

CRITICAL Before starting the experiment the participant must be trained on how to use the robotic device. To this aim, the participant will be asked to perform few movements in the synchronous condition with the eyes closed, at a frequency of approximately 1Hz, and to explore the whole back (Figure 4). Participants must be instructed not to count the time between pokes nor the number of pokes, but do the movement as learned during the training. The researcher can correct and guide the participants movement at this stage.

3. Prepare research subject for the robotic sensorimotor task (blindfold, headphones), robot should already be switched on and calibrated (see critical point above)
4. Provide instructions to participants (see critical point above)
5. Make the participant do few movements with the robotic devices, to make him/her understand how to control the robotic device, and to what resembles the touch on the back (habituation phase) applied by the back-robot.

[Data collection • Timing 15 minutes]

2. Perform the first sensorimotor stimulation (one sensorimotor condition pseudo-randomly chosen - e.g. synchronous).
 - Execute the function by typing the following into a matlab command-line terminal: "PH_Experiments". This will open a Guided User Interface (GUI; Figure 7) which is already populated with default values for a standard experiment. For example: a single trial using 0ms delay, for a total trial duration of 120 seconds, and one single trial (in "Nb trials").
 - Edit the default values so that they are consistent with the chosen condition.
 - Press "Launch" to start the experiment.
3. Remove the blindfold, and ask the person to fill in the questionnaire to assess the intensity of illusions (see Supplementary Table 1).
4. Take a short break to allow the participant to rest (usually several minutes, but can be adapted on the individual participant).
5. Perform the second sensorimotor stimulation (the other sensorimotor condition pseudo-randomly chosen - e.g. asynchronous).
 - In the GUI, which should still remain open, change the delay to the other delay (e.g. 500ms). The default values can still be changed according to the experiment's need.
 - Press "Launch" to start the experiment again.
6. Remove the blindfold, and ask the person to fill in the questionnaires to assess the intensity of illusions (see Supplementary Table 1).

Sensorimotor delay dependency

[Setup • Timing 5 minutes]

7. Prepare research subject for robotic sensorimotor task as described in Steps 3-5 above.

Appendix

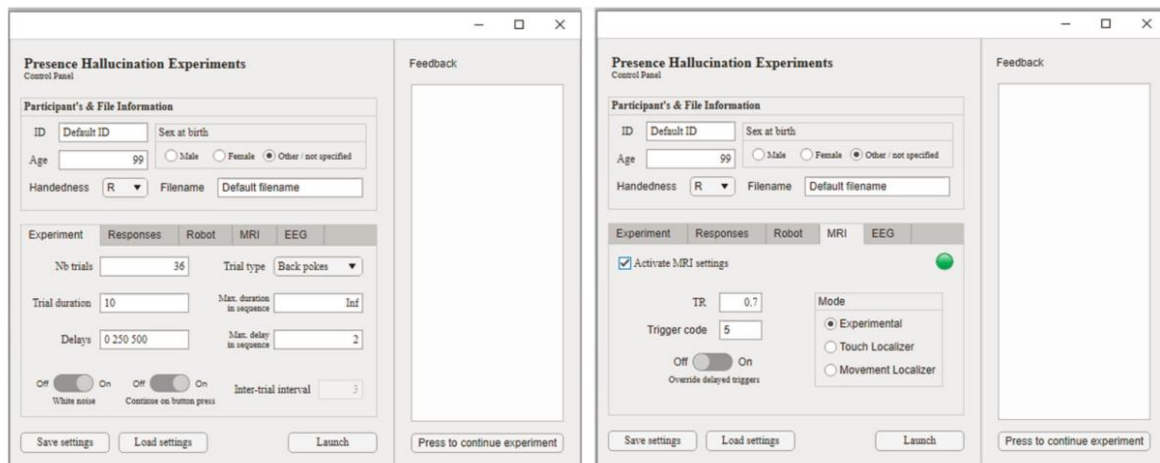


Fig. 7 | GUI for behavioral experiments. Left: view of the GUI with default values to conduct behavioral experiments. Each parameter can be adapted depending on the experimental needs. Positioning the cursor on the parameter icon will provide additional information on the role of the selected parameter. Right: GUI displaying the MRI settings. View of the GUI with default values to conduct behavioral experiments. Each parameter can be adapted depending on the experimental needs. Positioning the cursor on the parameter icon will provide additional information on the role of the selected parameter. By selecting 'Activate MRI settings' it will be possible to conduct MRI experiments.

[Data collection • Timing 20-25 minutes]

8. Perform sensorimotor stimulation (one sensorimotor condition pseudo-randomly chosen - e.g. synchronous).
 - Execute the function by typing the following into a matlab command-line terminal: "PH_Experiments". This will open a GUI (Supplementary Figure 7) which is already populated with default values for a standard experiment..
 - Update these values so that they appropriate for the current task. This can be done manually or by loading a file. E.g. Load the provided file "Settings_SensorimotorDelayDependency.m" using the button "Load settings" in the main page. This will prepare the GUI for this particular experiment. For example: delays randomly chosen amongst 0, 250 and 500ms, 36 repetitions in total, and a total duration of 10 seconds for each trial. These default values can also be changed depending on what is needed.

The duration of each experimental session will vary based on the number of delays, repetitions and duration of each trial.

9. Remove blindfold, and take a break for a few minutes.
10. Repeat sensorimotor stimulation
11. Remove blindfold, and do minutes break
12. Repeat sensorimotor stimulation
13. Remove blindfold

fMRI acquisition [Timing around 1h-1h10]

CRITICAL Before starting the experiment, the participant must be trained on how to use the robotic device. To this aim, the participant will be asked to perform a few movements in the synchronous condition with the eyes closed, at a frequency of ~1 Hz, and explore the whole space. The researcher

Appendix

can correct and guide the participant's movement at this stage. If the user's MRI data are BIDS compliant (<https://bids.neuroimaging.io/>), few modifications (e.g., modifying the scripts to recognize the user's functional files) will be necessary.

fMRI robotic setup and initial check-ups [Timing 10 mins]

1. Install the robotic system on the scanner bed, and connect the different cables with the two motors and the front-robot (Supplementary Fig. 2). Make sure to plug them in by properly following the labels.
2. Verify that the contact part of the back-robot is located at the center of the slide of the back-robot as well as the slider of the front-robot. The contact part of the back-robot has to have both 'legs parallel' (Supplementary Fig. 3). If it is not the case, manually change the front-robot slider to the middle, and for the back-robot, a recalibration would be needed in Step 10.
3. Verify that all screws are tight; if some are too loose, they need to be fixed. The same applies if the screws are too tight.

!CAUTION! Only use MR-compatible equipment to tighten or loosen screws in the robotic system; alternatively, take the robotic system out of the MR platform to do this step.

4. Open the code 'HapticGroup' and launch it to open the GUI interface using Visual Studio 2010 by pressing F5 + Ctrl. This will open the GUI interface, and an information window will appear saying either that 'joypad is connected' or that 'Joypad is disabled'; click 'OK'.
5. Initialize the robotic position by going to 'Initialization (I)', then 'Initialize NI card' and then 'Initialization (I)' and 'Zero-volt output'.
6. Turn on the control box.
7. Initialization of the devices: go on 'Experiment (E)', then 'Device1/2: initialization', and then click 'OK' without changing the parameters.
8. If the front and back robots were not in the correct position (from Step 4), change their position now:
 - The front robot can be moved to the middle by hand if not done previously
 - For the back-robot contact part, use the key '1' and '2' to move the first motor and the keys '3' and '4' to move the second motor. The contact part should be located in the middle of the slider (Supplementary Fig. 3)
 - When this is done, press 'enter' to save these initial positions and then 'Experiment (E)', 'Sampler OFF'
9. Select the input device, by going to 'Experiment (E)', 'Select input device (front-robot)' and then 'Device 1-Passive'.
10. Select the output device by going to 'Experiment (E)', 'Select output device (back-robot)' and then 'Device 2 -Position control'.

Appendix

The robotic system is ready to be started.

Participant preparation [Timing: 10-15 mins]

CRITICAL This orientation procedure is done before the experiment itself starts.

11. Complete and verify the MR safety questionnaire (the participant has to remove all metallic objects and is reminded of the magnetic field present in the MR).
12. Start the robotic device. On starting, you will see 'Experiment (E)', 'Sampler ON', and a window will appear to confirm the chosen input and output devices. Click 'Yes' if it is correct.
13. When the participant enters the MR scanner, explain the instructions for the task. Provide additional general information regarding the MR scanner (e.g., loud noise, earplugs necessary and emergency button).
14. Set the participant up on the mattress.
15. Ensure that the participant feels the back robot on their back and that the movement performed with the hand is done correctly.
16. Make sure that the participant is blindfolded, puts on the earplugs and headphones, and has access to the MR emergency button (which is placed on the chest).
17. Provide support to the elbows to reduce the amount that the head moves when the participant moves the robot. Tell them to limit their movement to a wrist movement to avoid moving the shoulder and hence the head.

fMRI acquisition [Timing: ~1h to 1h10]

18. When the participant enters the MR-scanner remind them of the instructions.
19. Execute the function by typing the following into a matlab command-line terminal: "PH_Experiments". This will open the GUI.
20. In the experiment tab, turn off the "Continue on button press", set the number of trials to 8, the trial duration to 30, the delays to "0 500", the trial type to "seconds", the "Inter-trial interval" to 18.
21. Go to the MRI tab. In the example shown in Figure 7, the repetition time was set to 2.5 seconds, and the MRI trigger was set to '5'. Modify the MRI trigger to the appropriate value for your configuration. The number of trials, trial duration, inter-trial interval, and the other experimental parameters will still follow the parameters input in the Experiment tab. Note however, that using MRI settings, will force the "Trial type" to be "Seconds". Change the default values as appropriate for the experiment.

Appendix

22. While the subject is performing the task make sure that the robotic system is moving properly and not blocked by inspecting visually in the Visual Studio interface that the “Slave Angle[rad]” is changing. <Troubleshooting>
23. Record the t1-weighted anatomical image.
24. Perform a second run of the robotic stimulation experiment.
 - remind the participant of the instructions.
 - If you haven’t closed the GUI, you may simply press “Launch” again, otherwise re-open the GUI by typing “PH_Experiments” in the command line as in step 9.<Troubleshooting>
25. Perform the “touch control task” experiment after reminding the instructions to the participant. In this task, an image is acquired while the person’s back is touched by the back robot.
 - While maintaining all experimental parameters, navigate to the MRI tab and select “Touch Localizer” under the mode panel.
 - Re-launch the experiment.<Troubleshooting>
26. Perform the “motor control task” experiment after reminding the instructions to the participant. In this task an image is acquired while the person pokes the front robot; with no other stimulus.
 - While maintaining all experimental parameters, navigate to the MRI tab and select “Motor Localizer” under the mode panel.
 - Re-launch the experiment.<Troubleshooting>
27. Switch the scanner off, and repeat the two experiments.
 - Using the software, choose “PH_Experiments” and change the “Trial duration” time to 30 sec and select 2 as number of trials.
 - Record the outcome using the behavioral questionnaire
28. On the Visual Studio interface, click on “Experiment(E)” then on “Sampler OFF”. When closing the program, a window will appear to reset the initial position of the back-robot, click on “Yes”.
29. Turn off the control box.
30. Take the subject out the MRI, remove the head coil, the blindfold, headphones and earplugs.

fMRI preprocessing [Timing: around 3-4h]

31. Once acquired, check that the functional and structural images are in BIDS format⁷⁷.

Appendix

CRITICAL This format is required, because if the user's MRI data are Brain Imaging Data Structure (BIDS) compliant (<https://bids.neuroimaging.io/>), few modifications (e.g., modifying the scripts to recognize the user's functional files) will be necessary.

32. Apply a standard preprocessing pipeline to the images, including the following steps:

- Correction of the functional images for slice timing
- Spatial realignment of the functional images
- Co-registration of the anatomical image with the mean functional image and segmentation of the anatomical image into grey matter, white matter and cerebro-spinal fluid (CSF) tissue.
- Normalisation to the Montreal Neurological Institute (MNI) brain template of both functional and anatomical images.
- Smoothing with a Gaussian kernel with full-width half-maximum of 6mm of the functional images.
- Assess head motion by performing a framewise displacement (FD) calculation 78.
- Use filters to remove low frequency confounds. The two experimental runs are filtered with a high-pass filter at 1/300 Hz, while the two control tasks are filtered with a high-pass filter at 1/100 Hz.

33. Activation contrasts: Apply a general linear model analysis to the functional images from both the experimental runs and the control tasks, for each subject. Model as regressors the following parameters:

- for the experimental runs, the two different conditions (i.e. synchronous and asynchronous) and in the control tasks, the robotic task (either corresponding to the participant's movement or to the robotic stimulation on the back)
- the auditory cues (lasting 1 sec).
- The six motion parameters for each run.

34. Also model as parametric modulators the total amount of movement performed during the conditions in which the participant is moving the front-robot (i.e. asynchronous, synchronous and motor task condition). If you wanted to do exactly what we did, then threshold the results with a threshold of $p < 0.001$ at voxel level and then cluster-wise FWE-corrected ($p < 0.05$) for multiple comparison.

Troubleshooting

Procedure 1

Step 1 (Front-robot (Geomagic) installation and initialization)

Serial connection error

Error description: "Serial connection ERROR message displayed when opening the controller application.

Appendix

Cause: The controller application was not able to find the COM port corresponding to the driver electronic, or the driver is not answering.

Solutions:

- a. Open windows devices manager under "port (COM & LPT)" check if the device "Silicon Labs CP210x..." is listed. If not, make sure the electric box is ON and the USB cable is plugged (if a USB hub is used make sure it is on). You can try changing USB port.
- b. If the device is listed make sure its driver is installed. If properly installed, the device should be listed as shown in the image below (except for COM number that may be different). If the icon is a yellow "!" or "?" that means the driver is not installed. The driver is in the installation folder [..\HapticGroup_tool_set_Backup\Installer\CP210x_driver\CP210xVCPInstaller_x64.exe].
- c. If you use Windows 7 or 8 it's possible that windows automatically installed a version of the driver that doesn't support a high baud rate. You can check the version in the device manager:
 - right click ->Property->driver tab. If the version is 10.X.X.X or above change it to 6.X.X.X
 - If not already done, install the CP210xVCPInstaller_x64.exe given in the installation folder. Then in the device manager: right clic->Property->driver tab->update driver -> Search driver on my computer-> chose among a list, then select version 6.X.X.X.
- d. If you use HapticGroup_V3_1 or lower version, the COM port of the electric box must have the lowest ID, so if other COM ports are used make sure the lowest ID is attributed to the electric driver (if not change it)

Front-robot (PHANToM Omni) cannot be opened

Error description: When enabling the sampler (during initialization or experiment) the controller application tells you it cannot connect to the master devices

Cause: The front-robot was not properly initialized, or you forgot to close the front-robot application

Solutions:

- a. Turn off and on the front-robot, wait for a few seconds and redo the initialization using the "Touch Setup" software. If you use the old version of the software don't forget to press the "apply" button before closing the application. Make sure you closed the Setup application before running the controller application.
- b. In some cases, if you have a USB version of the front-robot the controller application cannot connect to it (although the setup phase works fine). Currently, there is no available solution. You should use the Ethernet version of the front-robot or use a different computer

Step 2 (Back-robot installation and initialization)

Appendix

The robot doesn't move smoothly

Error description: When you do a smooth continuous trajectory with the front-robot, the motion of the pocking robot is “jumpy”. There are two possible causes.

Cause 1: The serial communication with the pocking robot has too much (and too variable) delay. This is because the communication is done with a USB virtual COM port, and if the same USB controller is shared among multiple peripherals the delay increases

Solution 1:

- a. If you are using a USB hub to connect both the front-robot and the robot (and if you have enough free USB ports directly on the computer) try removing the hub and connect both directly to the computer. If you have a laptop with one USB-3 port (the blue one), use this port for the robot (alternatively if you use a USB-3 hub for both the robot and front-robot, connect the hub to this port). If you have other USB peripherals connected to the computer remove them.
- b. In some cases, this problem is more pronounced when the computer is in low power mode. The low power mode is usually the default mode on a laptop if it's running on battery, thus make sure your laptop is connected to its power supply.

Cause 2: This is less likely, but if the computer is performing some heavy tasks (like software installation, or antivirus scan), the control loop cannot be run with 2ms sample time.

Solution 2: Kill those tasks while you are using the robot.

Motion is different between front and back- and robot and/or bad virtual wall orientation

Error description: The virtual wall is not in a good position (not vertical or not flat), and/or the displacements of the pocking robot don't match the front-robot (even if you are not touching the boundary of the workspace). For example, a straight motion of the input corresponds to a curved one of the outputs, or a horizontal one corresponds to a “not horizontal” one.

Cause: the front- or back-robot hasn't been properly calibrated

Solution: Go back to the initialization mode, plug (for 3 seconds) and unplug the pen of the front-robot (you should hear a "click"). Then reset the initial position of the complete system by redoing the calibration (press the "i" key).

The relative position between front- and back-robot changes (drifts) during the experiment

Error description: After the initialization, the motion of the pocking robot corresponds to the motion of the front-robot (for example a straight-line trajectory of the input correspond to the same straight line of the output), but during the experiment the relative position change (the same straight

Appendix

trajectory of the input correspond to a different trajectory of the output). There are two possible causes.

Cause 1: The position of the front-robot is drifting during the experiment because one of its position sensors is damaged.

Diagnosis of the problem: To make sure this is the problem, close everything, open the front-robot application, and redo the initialization/calibration of the front-robot. Once it's done stay in the front-robot application (where you have the 3D representation of the device) and do some motion. You should see that the position of the 3D model and the position of the real device are drifting apart

Solution 1: No solutions, you need to change the front-robot

Causes 2: The position of the robot is drifting during the experiment because one of the couplings (of axis 1 or 3) is not properly tightened.

Diagnosis of the problem: Initialise the system (the same way you do usually). Then before starting the experiment set the motor current of axis 1 and 3 to the max value (5.5 amperes) and click apply the change (do not click "overwrite default", or if you don't forget to change back the value next time you use the system). Once the experiment is running, do not touch the front-robot, and, by hand, force changes the position of the pocking robot (front-back and vertical motion, no needs to move it sideways). If the robot doesn't come back to its initial position when you release it that means one of the couplings is not tight

Solution 2: Stop the experiment and tight the corresponding screw

Procedure 2, Step 25, 26,27.

The back-robot does not move anymore.

It is possible that the motors stop moving properly during the experiment. When this happens, go to "Experiment (E)", "Sampler OFF" and then turn off the controller (motor driver). Wait a few seconds and turn them on again before starting the initialization again: "Experiment (E), Sampler ON".

Anticipated results

Behavior

Spatio-temporal sensorimotor stimulation

Based on our previous results ^{31,32,42} we predict that in the condition of asynchronous sensorimotor stimulation a participant will experience a stronger riPH and stronger passivity experience (feeling of being touch by someone else), in both the behavioral (Figure 2A & 2C) and the MRI settings (Figure 2B). Concerning the implicit measures of riPH, during the asynchronous stimulation, participants may also show a drift in self-location in the backward direction and indicate, and have a tendency to

Appendix

overestimate the number of individuals close-by (numerosity estimation) ²⁹. In the synchronous sensorimotor stimulation, we predict that the participants will experience a stronger sensation of touching themselves (self-touch), despite extending their arms in front of their bodies. The synchronous sensorimotor stimulation and stimulation with force feedback are further associated with a drift of the subject's body toward the front position and lower estimation of individuals close-by (when compared to the asynchronous condition) ²⁹. Generally, the absence of a force feedback results in stronger riPH in the asynchronous condition. Based on our previous findings ³¹, we expect that patients with Parkinson's disease and PH are more sensitive (higher ratings in the questionnaires) to riPH, and in some of these patients the phenomenology of the riPH resembles to what experience in daily life.

Sensorimotor delay dependency

Based on our previous results ³¹, we expect that the probability to report/experience a riPH increases with the increase of the sensorimotor conflict (i.e. the longer the delay between movement and touch on the back) (see Figure 2C). In addition, we expect that patients with Parkinson's disease and PH (vs. patients without PH) show a higher sensitivity to the riPH already with 0ms delay (as indicated by the higher intercept), and a stronger sensitivity to the delays (as indicated by the steeper slope) ³¹.

fMRI

Based on our previous findings ³¹, we expect that the contrast asynchronous > synchronous (i.e. showing brain regions more activated in the asynchronous condition compared to the synchronous condition) will activate the right pSTS, the inferior frontal gyrus (IFG), the right anterior insula and the right medial prefrontal cortex (mPFC).

Regarding the control tasks, we expect that the action of moving the front robot (motor control task) will result in activations in a large cortical-subcortical network including contralateral sensorimotor cortex (depending on which hand the participant uses during the robotic stimulation), putamen and thalamus and bilateral premotor cortex, SMA, parietal operculum (secondary somatosensory cortex (SII), supramarginal gyrus (SMG)), and the ipsilateral cerebellum. We expect that the experience of being touch on the back (touch control task) will activate the bilateral parietal operculum including SMG, SII and the superior temporal gyrus (STG).

We also expect based on our previous findings ³¹ that the conjunction analysis (asynchronous > [motor + touch] \cap synchronous > [motor + touch]) will reveal a subcortical-cortical network in the left sensorimotor cortex, bilateral supplementary motor area, right inferior parietal cortex, left putamen, and right cerebellum. This contrast highlights the brain regions involved in the spatial sensorimotor conflict between the sensorimotor movement of the hand in front space and the feedback in the back, independent of the sensorimotor delay.

For the subjective questionnaire assessing the PH (in the dataset example assessed at the end of the scanning session), we predict that in the condition of asynchronous sensorimotor stimulation a participant will experience a stronger riPH and stronger passivity experience (feeling of being touch by someone else).

Appendix

Although the MRI robotic approach has not been yet applied to Parkinson's patients, based on our previous findings ³¹, we expect that patients with Parkinson's disease and PH would have higher activity in the pSTS and IFG compared to the patients with Parkinson's disease without PH in the asynchronous versus synchronous condition since both of these regions were identified as part of the PH-network (for more details see ³¹). We would expect the same in dementia with Lewy bodies with PH compared to patients without PH. Regarding the control tasks, we would expect no difference in activations between the patients (Parkinson's disease or DLB) with or without PH.

Data availability

MRI data are available on zenodo.org (<https://zenodo.org/record/4423384#.YkKyHDWxVmN>), behavioral data can be found on gitlab (<https://gitlab.epfl.ch/fbernasc/np-p210507a.git>)

Code availability

The codes to control the robots are uploaded in gitlab <https://gitlab.epfl.ch/fbernasc/roboticsph.git>

References

1. Tracy, D. K. & Shergill, S. S. Mechanisms Underlying Auditory Hallucinations-Understanding Perception without Stimulus. *Brain Sci.* 3, 642–669 (2013).
2. Corlett, P. R. et al. Hallucinations and Strong Priors. *Trends Cogn. Sci.* 23, 114–127 (2019).
3. Larøi, F. et al. An epidemiological study on the prevalence of hallucinations in a general-population sample: Effects of age and sensory modality. *Psychiatry Res.* 272, 707–714 (2019).
4. Badcock, J. C., Dehon, H. & Larøi, F. Hallucinations in Healthy Older Adults: An Overview of the Literature and Perspectives for Future Research. *Front. Psychol.* 8, (2017).
5. Badcock, J. C. et al. Hallucinations in Older Adults: A Practical Review. *Schizophr. Bull.* 46, 1382–1395 (2020).
6. Ohayon, M. M. Prevalence of hallucinations and their pathological associations in the general population. *Psychiatry Res.* 97, 153–164 (2000).
7. Collerton, D., Perry, E. & McKeith, I. Why people see things that are not there: a novel Perception and Attention Deficit model for recurrent complex visual hallucinations. *Behav. Brain Sci.* 28, 737–757; discussion 757-794 (2005).
8. Insel, T. R. Rethinking schizophrenia. *Nature* 468, 187–193 (2010).
9. Arnaoutoglou, N. A., O'Brien, J. T. & Underwood, B. R. Dementia with Lewy bodies — from scientific knowledge to clinical insights. *Nat. Rev. Neurol.* 15, 103–112 (2019).
10. Ffytche, D. H. et al. The psychosis spectrum in Parkinson disease. *Nat. Rev. Neurol.* 13, 81–95 (2017).
11. Millan, M. J. et al. Altering the course of schizophrenia: progress and perspectives. *Nat. Rev. Drug Discov.* 15, 485–515 (2016).
12. Llorca, P. M. et al. Hallucinations in schizophrenia and Parkinson's disease: an analysis of sensory modalities involved and the repercussion on patients. *Sci. Rep.* 6, 38152 (2016).
13. Jaspers, K. Über leibhaftige Bewußtheiten (Bewußtheitstäuschungen), ein psychopathologisches Elementarsymptom. *Z Pathopsychol.* 2. 150-161 (1913).
14. McKeith, I. G. et al. Diagnosis and management of dementia with Lewy bodies: Fourth consensus report of the DLB Consortium. *Neurology* 89, 88–100 (2017).
15. Buracchio, T., Arvanitakis, Z. & Gorbien, M. Dementia with Lewy bodies: current concepts. *Dement. Geriatr. Cogn. Disord.* 20, 306–320 (2005).
16. Walker, Z., Possin, K. L., Boeve, B. F. & Aarsland, D. Lewy body dementias. *Lancet Lond. Engl.* 386, 1683–1697 (2015).
17. Nagahama, Y. et al. Classification of psychotic symptoms in dementia with Lewy bodies. *Am. J. Geriatr. Psychiatry Off. J. Am. Assoc. Geriatr. Psychiatry* 15, 961–967 (2007).
18. Eversfield, C. L. & Orton, L. D. Auditory and visual hallucination prevalence in Parkinson's disease and dementia with Lewy bodies: a systematic review and meta-analysis. *Psychol. Med.* 49, 2342–2353 (2019).
19. Nicastro, N., Eger, A. F., Assal, F. & Garibotto, V. Feeling of presence in dementia with Lewy bodies is related to reduced left frontoparietal metabolism. *Brain Imaging Behav.* 14, 1199–1207 (2020).
20. Goetz, C. G., Fan, W., Leurgans, S., Bernard, B. & Stebbins, G. T. The malignant course of 'benign hallucinations' in Parkinson disease. *Arch. Neurol.* 63, 713–716 (2006).

Appendix

21. Fénelon, G., Soulas, T., Langavant, L. C. de, Trinkler, I. & Bachoud-Lévi, A.-C. Feeling of presence in Parkinson's disease. *J Neurol Neurosurg Psychiatry* 82, 1219–1224 (2011).
22. Lenka, A., Pagonabarraga, J., Pal, P. K., Bejr-Kasem, H. & Kulisvesky, J. Minor hallucinations in Parkinson disease: A subtle symptom with major clinical implications. *Neurology* (2019) doi:10.1212/WNL.0000000000007913.
23. Ravina, B. et al. Diagnostic criteria for psychosis in Parkinson's disease: report of an NINDS, NIMH work group. *Mov. Disord. Off. J. Mov. Disord. Soc.* 22, 1061–1068 (2007).
24. Pagonabarraga, J. et al. Neural correlates of minor hallucinations in non-demented patients with Parkinson's disease. *Parkinsonism Relat. Disord.* 20, 290–296 (2014).
25. Bejr-Kasem, H. et al. Minor hallucinations reflect early gray matter loss and predict subjective cognitive decline in Parkinson's disease. *Eur. J. Neurol.* 28, 438–447 (2021).
26. Rosenthal, R. & Fode, K. L. Psychology of the Scientist: V. Three Experiments in Experimenter Bias. *Psychol. Rep.* 12, 491–511 (1963).
27. Adler, N. E. Impact of Prior Sets Given Experimenters and Subjects on the Experimenter Expectancy Effect. *Sociometry* 36, 113–126 (1973).
28. Rogers, S., Keogh, R. & Pearson, J. Hallucinations on demand: the utility of experimentally induced phenomena in hallucination research. *Philos. Trans. R. Soc. B Biol. Sci.* 376, 20200233 (2021).
29. Blanke, O. et al. Neurological and robot-controlled induction of an apparition. *Curr. Biol. CB* 24, 2681–2686 (2014).
30. Hara, M. et al. A novel manipulation method of human body ownership using an fMRI-compatible master–slave system. *J. Neurosci. Methods* 235, 25–34 (2014).
31. Bernasconi, F. et al. Robot-induced hallucinations in Parkinson's disease depend on altered sensorimotor processing in fronto-temporal network. *Sci. Transl. Med.* 13, (2021).
32. Salomon, R. et al. Sensorimotor Induction of Auditory Misattribution in Early Psychosis. *Schizophr. Bull.* (2020) doi:10.1093/schbul/sbz136.
33. Ford, J. M. & Mathalon, D. H. Electrophysiological evidence of corollary discharge dysfunction in schizophrenia during talking and thinking. *J. Psychiatr. Res.* 38, 37–46 (2004).
35. Heinks-Maldonado, T. H. et al. Relationship of imprecise corollary discharge in schizophrenia to auditory hallucinations. *Arch. Gen. Psychiatry* 64, 286–296 (2007).
36. Brugger, P., Regard, M. & Landis, T. Unilaterally Felt 'Presences': The Neuropsychiatry of One's Invisible Doppelgänger. *Cogn. Behav. Neurol.* 9, (1996).
37. Blanke, O., Ortigue, S., Coeytaux, A., Martory, M.-D. & Landis, T. Hearing of a presence. *Neurocase* 9, 329–339 (2003).
38. Arzy, S., Seeck, M., Ortigue, S., Spinelli, L. & Blanke, O. Induction of an illusory shadow person. *Nature* 443, 287 (2006).
39. Blakemore, S. J., Wolpert, D. & Frith, C. Why can't you tickle yourself? *Neuroreport* 11, R11-16 (2000).
40. Ehrsson, H. H., Holmes, N. P. & Passingham, R. E. Touching a rubber hand: feeling of body ownership is associated with activity in multisensory brain areas. *J. Neurosci. Off. J. Soc. Neurosci.* 25, 10564–10573 (2005).
41. Ionta, S. et al. Multisensory mechanisms in temporo-parietal cortex support self-location and first-person perspective. *Neuron* 70, 363–374 (2011).

Appendix

42. Hara, M. et al. A novel approach to the manipulation of body-parts ownership using a bilateral master-slave system. in 2011 IEEE/RSJ International Conference on Intelligent Robots and Systems 4664–4669 (2011). doi:10.1109/IROS.2011.6094879.
43. Serino, A. et al. Thought consciousness and source monitoring depend on robotically controlled sensorimotor conflicts and illusory states. *iScience* 24, 101955 (2021).
44. Weiskrantz, L., Elliott, J. & Darlington, C. Preliminary observations on tickling oneself. *Nature* 230, 598–599 (1971).
45. Pozeg, P., Rognini, G., Salomon, R. & Blanke, O. Crossing the Hands Increases Illusory Self-Touch. *PLoS ONE* 9, (2014).
46. Blanke, O., Slater, M. & Serino, A. Behavioral, Neural, and Computational Principles of Bodily Self-Consciousness. *Neuron* 88, 145–166 (2015).
47. Park, H.-D. & Blanke, O. Coupling Inner and Outer Body for Self-Consciousness. *Trends Cogn. Sci.* 23, 377–388 (2019).
48. Kahn, R. S. et al. Schizophrenia. *Nat. Rev. Dis. Primer* 1, 1–23 (2015).
49. Fletcher, P. C. & Frith, C. D. Perceiving is believing: a Bayesian approach to explaining the positive symptoms of schizophrenia. *Nat. Rev. Neurosci.* 10, 48–58 (2009).
50. Orepic, P., Rognini, G., Kannape, O. A., Faivre, N. & Blanke, O. Sensorimotor conflicts induce somatic passivity and louden quiet voices in healthy listeners. *Schizophr. Res.* 231, 170–177 (2021).
51. Stripeikyte, G. et al. Fronto-Temporal Disconnection Within the Presence Hallucination Network in Psychotic Patients With Passivity Experiences. *Schizophr. Bull.* sbab031 (2021) doi:10.1093/schbul/sbab031.
52. Jones, S. A. V. & O'Brien, J. T. The prevalence and incidence of dementia with Lewy bodies: a systematic review of population and clinical studies. *Psychol. Med.* 44, 673–683 (2014).
53. Nicolas NICASTRO, STRIPEIKYTE, G., ASSAL, F., GARIBOTTO, V. & Blanke, O. Premotor and fronto-striatal mechanisms associated with presence hallucinations in dementia with Lewy bodies. *Neuroimage Clin.* In Press,.
54. Soulas, T., Cleret de Langavant, L., Monod, V. & Fénelon, G. The prevalence and characteristics of hallucinations, delusions and minor phenomena in a non-demented population sample aged 60 years and over. *Int. J. Geriatr. Psychiatry* 31, 1322–1328 (2016).
55. O'Callaghan, C. et al. Impaired sensory evidence accumulation and network function in Lewy body dementia. *Brain Commun.* (2021) doi:10.1093/braincomms/fcab089.
56. Schumacher, J. et al. Functional connectivity in mild cognitive impairment with Lewy bodies. *J. Neurol.* (2021) doi:10.1007/s00415-021-10580-z.
57. Allefeld, C., Pütz, P., Kastner, K. & Wackermann, J. Flicker-light induced visual phenomena: frequency dependence and specificity of whole percepts and percept features. *Conscious. Cogn.* 20, 1344–1362 (2011).
58. Pearson, J. et al. Sensory dynamics of visual hallucinations in the normal population. *eLife* 5, e17072 (2016).
59. Wackermann, J., Pütz, P. & Allefeld, C. Ganzfeld-induced hallucinatory experience, its phenomenology and cerebral electrophysiology. *Cortex J. Devoted Study Nerv. Syst. Behav.* 44, 1364–1378 (2008).

Appendix

60. Mason, O. J. & Brady, F. The psychotomimetic effects of short-term sensory deprivation. *J. Nerv. Ment. Dis.* 197, 783–785 (2009).
61. Merabet, L. B. et al. Visual hallucinations during prolonged blindfolding in sighted subjects. *J. Neuro-Ophthalmol. Off. J. North Am. Neuro-Ophthalmol. Soc.* 24, 109–113 (2004).
62. Zarkali, A., Lees, A. J. & Weil, R. S. Flickering Stimuli Do Not Reliably Induce Visual Hallucinations in Parkinson's Disease. *J. Park. Dis.* 9, 631–635 (2019).
63. Ellson, D. G. Hallucinations produced by sensory conditioning. *J. Exp. Psychol.* 28, 1–20 (1941).
64. Powers, A. R., Mathys, C. & Corlett, P. R. Pavlovian conditioning–induced hallucinations result from overweighting of perceptual priors. *Science* 357, 596–600 (2017).
65. Vollenweider, F. X. & Preller, K. H. Psychedelic drugs: neurobiology and potential for treatment of psychiatric disorders. *Nat. Rev. Neurosci.* 21, 611–624 (2020).
66. Carhart-Harris, R. L. How do psychedelics work? *Curr. Opin. Psychiatry* 32, 16–21 (2019).
67. Parvizi, J. et al. Altered sense of self during seizures in the posteromedial cortex. *Proc. Natl. Acad. Sci. U. S. A.* 118, e2100522118 (2021).
68. Penfield, W. & Perot, P. THE BRAIN'S RECORD OF AUDITORY AND VISUAL EXPERIENCE. A FINAL SUMMARY AND DISCUSSION. *Brain J. Neurol.* 86, 595–696 (1963).
69. Heydrich, L., Lopez, C., Seeck, M. & Blanke, O. Partial and full own-body illusions of epileptic origin in a child with right temporoparietal epilepsy. *Epilepsy Behav.* EB 20, 583–586 (2011).
70. Blanke, O., Perrig, S., Thut, G., Landis, T. & Seeck, M. Simple and complex vestibular responses induced by electrical cortical stimulation of the parietal cortex in humans. *J. Neurol. Neurosurg. Psychiatry* 69, 553–556 (2000).
71. PENFIELD, W. & BOLDREY, E. SOMATIC MOTOR AND SENSORY REPRESENTATION IN THE CEREBRAL CORTEX OF MAN AS STUDIED BY ELECTRICAL STIMULATION¹. *Brain* 60, 389–443 (1937).
72. Frith, C. D. & Done, D. J. Experiences of alien control in schizophrenia reflect a disorder in the central monitoring of action. *Psychol. Med.* 19, 359–363 (1989).
73. Tsakiris, M., Hesse, M. D., Boy, C., Haggard, P. & Fink, G. R. Neural signatures of body ownership: a sensory network for bodily self-consciousness. *Cereb. Cortex N. Y. N* 1991 17, 2235–2244 (2007).
74. Ehrsson, H. H., Spence, C. & Passingham, R. E. That's my hand! Activity in premotor cortex reflects feeling of ownership of a limb. *Science* 305, 875–877 (2004).
75. Lenggenhager, B., Tadi, T., Metzinger, T. & Blanke, O. Video Ergo Sum: Manipulating Bodily Self-Consciousness. *Science* 317, 1096–1099 (2007).
76. Farrer, C., Valentin, G. & Hupé, J. M. The time windows of the sense of agency. *Conscious. Cogn.* 22, 1431–1441 (2013).
77. Blakemore, S.-J., Wolpert, D. M. & Frith, C. D. Central cancellation of self-produced tickle sensation. *Nat. Neurosci.* 1, 635–640 (1998).
78. Gorgolewski, K. J. et al. The brain imaging data structure, a format for organizing and describing outputs of neuroimaging experiments. *Sci. Data* 3, 160044 (2016).
79. Power, J. D., Barnes, K. A., Snyder, A. Z., Schlaggar, B. L. & Petersen, S. E. Spurious but systematic correlations in functional connectivity MRI networks arise from subject motion. *Neuroimage* 59, 2142–2154 (2012).

Appendix

Curriculum Vitae



HERBERTO DHANIS

PhD Candidate

✉ herberto.dhanis@gmail.com 📍 Chemin des Mines 9, Geneva, Switzerland

Portuguese: Native | English: Proficient | French: B2 | Spanish: A2

SKILLS

Hard

Signal processing
Statistic and statistical models
Functional brain analysis
Temporal models
Scientific writing
Project planning and management

Professional

Communication
Leadership
Independence
Logical thinking
Problem solving + fast learning
Oral presentation

Programming

Proficient: MATLAB, R, python
Intermediate: Java
Basis in: C++, C#, HTML, CSS, Javascript

OUTREACH & AWARENESS

"Long Covid on the short"

Creator of a YouTube channel dedicated to scientific communication on research behind Long Covid pathology
11/2022 - to present

3rd Finalist in My Thesis in 180 seconds | EPFL 2022 edition
Scientific communication competition

Gambling addiction: talks, films and exhibitions

Bringing mathematicians, psychologists and former addicts together, to raise awareness on gambling addiction
09/2010 - 01/2011

PROFESSIONAL EXPERIENCE

Swiss Federal Institute of Technology (EPFL) | PhD Candidate

10/2017 - 02/2023

- Implemented a cross-platform project requiring programming skills in MATLAB, R and python
- Managing clinical imaging and software for a large multi-center project
- Functional magnetic resonance imaging. Skills in SPM, CONN, ART, CAPs, and dynamic temporal modelling
- Study design and execution, formal modelling analysis, interpretation, scientific communication and writing, teaching

Affinity | Junior Java Developer

04/2017 - 08/2017

- Development of an API (Java 8, Spring, Hibernate) for an online portal
- Recognized as a valid asset due to efficiency and productivity

NOVA University of Lisbon | Research Assistant

01/2014 - 07/2015

- Identifying neural markers of sleep stages using EEG
- Development of an automatic sleep classifier with MATLAB

GLINTT | Complex Deals Consultant Trainee

08/2014 - 09/2014

- Advising on interoperability of medical devices
- Market analysis of health information systems

Pedro Pharmacy | Back Office Assistant

06/2013 - 07/2013

- Development of a program to assist in dosage and preparation of patient-tailored medications

EDUCATION

PhD in Neuroscience

Swiss Federal Institute of Technology (EPFL)

10/2017 - 02/2023

Neuro-engineering of hallucinations through fMRI neurofeedback: Understanding and modulating neural dynamics for the control of hallucinations in health and disease

MSc Thesis | Exchange within Erasmus +

Swiss Federal Institute of Technology (EPFL)

02/2016 - 10/2016

Bio-engineering and Innovation in Neurosciences | Erasmus +
Université Paris-Descartes & ParisTech Group

09/2015 - 10/2016

BSc and MSc in Biomedical Engineering

NOVA University of Lisbon

02/2016 - 10/2016

Bachelor: 16/20 | Master 18/20 | Master Thesis: 19/20

CONFERENCES

TALK | Organisation for Human Brain Mapping (OHBM) 2022

10-minute oral session: CAP-based fMRI neurofeedback enables bidirectional regulation of induced hallucination networks

TALK | Alpine Brain Imaging Meeting (ABIM) 2022

15-minute oral session: CAP-based fMRI neurofeedback enables bidirectional regulation of induced hallucination networks

POSTER | ABIM 2020 and OHBM 2020

Control over brain networks associated with robotically-mediated hallucinations through dynamic functional connectivity-based real-time fMRI neurofeedback

POSTER | Real-time fMRI neurofeedback conference (rtFIN) 2019

Control over robotically-mediated hallucinations through dynamic functional connectivity real-time fMRI neurofeedback

POSTER | ABIM 2019

Dynamic functional connectivity of the robotically-induced Presence Hallucination

TEACHING

Neuroscience for Engineers | 2018 – 2019 – 2020

Image Processing I | 2019

Neuroscience III | 2020

PUBLICATIONS

Dhanis, et al. (2022) *NeuroImage*. Robotically-induced hallucinations triggers subtle changes in brain network dynamics

Dhanis, et al. (in preparation) Sustained bidirectional self-regulation of hallucination networks through real-time fMRI neurofeedback

Dhanis, et al. (in preparation) Robot-induced hallucinations and PLS-CAPs reveal neural correlates of preserved cognitive and psychosis-like traits

Dhanis, et al. (in preparation) Mega-analysis of presence hallucination induction experiments using robotically mediated sensorimotor conflicts

Bolton T., (...) Dhanis, H., (...), et al. TbCAPs: a toolbox for co-activation pattern analysis. *NeuroImage*, 211, 116621

Serino, A., (...) Dhanis, H., (...), et al. Thought consciousness and source monitoring depend on robotically controlled sensorimotor conflicts and illusory states. *iScience*, 24(1), 101955

Potheegadoo, J., (...) Dhanis, H., (...), et al. Presence Hallucinations during locomotion in patients with Parkinson's Disease. *Movement Disorders Clinical practice*, 24(1), 101955

Bernasconi, F., (...) Dhanis, H., (...), et al. Neuroscience robotics for controlled induction and real-time assessment of hallucinations. *Nature Protocols*, 24(1), 101955

Curriculum Vitae

Curriculum Vitae
