ALIUS BULLETIN

exploring the diversity of consciousness

Martin Fortier, Daniel Friedman, Matthieu Koroma & Raphaël Millière (eds.)

n°4
DECEMBER 2020
aliusresearch.org

Acknowledgments

The editors wish to express their gratitude to ALIUS members for their valuable help in proofreading the interviews, and to all the contributors for accepting to participate to this issue.

Foreword

About ALIUS

ALIUS is an international and interdisciplinary research group dedicated to the investigation of all aspects of consciousness, with a specific focus on non- ordinary or understudied conscious states traditionally classified as *altered states of consciousness*.

In Latin, *alius* means “different”. This lexical choice reflects the group’s mission to study the diversity of consciousness in a systematic manner. ALIUS puts a particular stress on the need for a naturalistic approach to all aspects of consciousness, including states and experiences which have long been unduly associated to parapsychology and pseudoscientific hypotheses.

To this end, it fosters a unique interdisciplinary collaboration of researchers, involving neuroscientists, psychologists, philosophers of mind, psychiatrists and anthropologists, towards the development of a systematic and scientific model of consciousness supported by both theoretical work and experimental studies. This collaboration may take the form of joint articles, blog posts, editorial work on special issues, thematic workshops and international conferences.

Find out more about the group on the website: aliusresearch.org `

About the Bulletin

The ALIUS Bulletin is an annual publication featuring in-depth interviews with prominent scholars working on consciousness and its altered states (ASCs). The goal of the Bulletin is to present a clear outline of current research on ASCs across a variety of disciplines, with an emphasis on empirical work. It also aims at dispelling the widespread stigma that still plagues the notion of ASC, while allowing a wider audience to discover rigorous scientific work on the topic presented by authors in their own words.

Table of contents

Tribute to Martin Fortier

by ALIUS Research Group

Cite as: ALIUS Research Group (2020). Tribute to Martin Fortier by ALIUS Research Group. *ALIUS Bulletin*, *4*, 1-14, https://doi.org/10.34700/aqmn-r841

ALIUS Research Group
aliusresearchgroup@gmail.com

Abstract

Martin Fortier has been a wonderful friend and a brilliant colleague, gone far too soon to realize his extraordinary potential despite his many precocious achievements. More than anyone, he embodied the spirit of interdisciplinarity. His passion for the scientific study of consciousness was inspiring, altering the trajectory of many researchers in our network and beyond. This text is intended to provide a faithful, up-to-date, and accurate representation of Martin's wide-ranging research by adapting in large part from Martin's own summary of his work written about a year before his passing.

keywords: *hallucinogens, bayesian constructivism, neuropharmacology, psychedelics, religion, consciousness, homo neurochemicus, HUTHAC*

On April 11th, 2020, Martin Fortier tragically passed away after a long and harrowing battle with cancer. He was thirty years old. We mourn the loss of a wonderful friend and a brilliant colleague, gone far too soon to realize his extraordinary potential despite his many precocious achievements.

Martin co-founded ALIUS in 2016, and has been a pillar of our group since its inception. More than anyone, he embodied the spirit of interdisciplinarity at the core of ALIUS' mission. His unique academic trajectory reveals in that respect: after a Bachelor's degree in philosophy at the Sorbonne, Martin obtained two Master's degrees in parallel - one in philosophy, and another in anthropology. He then started a Ph.D. in cognitive science at the Institut Jean Nicod, under the supervision of Jerôme Dokic. His dissertation, which remains unfinished, is an impressive *tour de force* seamlessly connecting ideas and data from philosophy, anthropology, psychology, neuroscience, and pharmacology.

Martin's ideal of interdisciplinarity was rooted in his aversion for empirical blindspots, partial evidence, and culturally-grounded intuitions. He encouraged philosophers to seek the support of empirical data and refrain from making underspecified or speculative claims about topics that can be empirically determined. He invited neuroscientists and psychologists to be more aware of the narrow cultural background of their experiments, that often recruit "WEIRD" participants - those raised in a society that is Western, Educated, Industrialized, Rich, and Democratic. Finally, he incited anthropologists to relate their theories to empirically-grounded hypotheses about the cognitive architecture of the human mind.

Martin strived to apply these injunctions to his own research. His life and work embodied many intellectual virtues rarely conjoined in a single individual to such a high degree: curiosity, passion, meticulousness, rigor, tenacity. Those who knew him remember how he could rekindle one's interest in research or spur ideas for a whole new project within a single conversation - so infectious was his love of learning and his enthusiasm for charting the human mind. When he set out to study a topic, he would get consumed by reading about it and leave no stone unturned. In this quest for encyclopedic knowledge, he devoured articles and books from many disciplines, and had an impressive capacity for elegantly synthesizing them in order to yield new insights. His research combined a number of methods and tools borrowed from his interdisciplinary expertise: conceptual analysis from analytic philosophy; theorization based on existing anthropological, psychological, neuroscientific and biological data; collection of new experimental data, and collection of ethnographic data. His fieldwork career included work with Shipibo-Konibo communities of the Ucayali region in the Peruvian Amazon, as well as Huni Kuin (Cashinahua) communities on the Peruvian side of the Purus region.

Martin liked to organize his many research interests within two broad categories. In the first category was his research on the structure of abstract knowledge in indigenous Amazonian cultures; in the second, his research on hallucinogens, culture, and neurobiology.

One of the key ideas behind Martin's work on abstract knowledge in Amazonian cultures is Bayesian constructivism. There is a long-lasting controversy in psychology between nativists, who think abstract knowledge is innately present in humans, and empiricists, who believe it is simply abstracted from regularities detected by the senses, or socially constructed. Martin favored a third proposal - Bayesian constructivism - according to which abstract knowledge is more than the simple accumulation of sensory data, but is nonetheless gradually constructed through development rather than being innately given or directly taught. One implication of this view is that different abstract structures may be constructed given different kinds of experience. Martin's work explored the fruitfulness of the Bayesian constructivist framework in studying how Amazonian indigenous people categorize and reason about the world.

One of Martin's targets was psychological essentialism, the view that external properties are explained by inner ones, and that identity is defined by inner and fixed properties rather than superficial and malleable ones. Many authors claim that essentialism is found in all cultures across the world. Martin's work challenged this claim. In particular, he argued that Cashinahua people use non-essentialist forms of reasoning. For them, many inner properties are explained by external ones and are thus highly malleable. He claimed that Cashinahua non-essentialism obtains both in the domains of ethnozoology (reasoning about different species of animals) and ethnosociology (reasoning about different ethnic groups).

Martin further considered the implications of the existence of non-essentialist reasoning in Amazonian cultures for anthropology. Anthropologists such as Descola and Viveiros de Castro have proposed that what is special in the way Amazonian people think is that they ascribe mental and bodily properties differently than we do (for example, as compared to us, they "over-ascribe" mental properties to animals and plants). Martin's work on non-essentialism suggests instead that what differentiates Amazonian thinking from our own thinking is not so much how mental or bodily properties are ascribed to beings of the world, but rather whether these properties are deemed fixed or malleable. On his view, we tend to

conceptualize many inner properties as being fixed while indigenous Amazonian people conceptualize them as being highly malleable. In ethnobiology, essentialist thinking has often been associated with taxonomic thinking. Several authors have maintained that ethnobiological knowledge is universally organized in taxonomic trees. In line with the Bayesian proposal, Martin explored the hypothesis that a given domain can be organized around various non-essentialist structures. His preliminary findings suggested that in some Amazonian cultures, abstract knowledge regarding biological beings is organized around complex and dynamical networks rather than taxonomic trees.

Martin also investigated color categorization in Amazonian cultures. Some authors claim that color categories are organized according to universal principles in every culture around the world. Martin's fieldwork with the Shipibo-Konibo brought to light five basic color terms that work consistently with the color categories described in the World Color Survey. However, it also revealed the existence of a myriad of ad hoc color terms. For example, to refer to red, the Shipibo-Konibo often use the word "*jimi*" ("blood") even though they have a basic color term specifically referring to red ("*joshin*"). The ad hoc color terms used by the Shipibo-Konibo are polysemous (i.e., they can be used to refer to different colors if the original anchor object features different colors) and if speakers are asked about their meaning, they will first mention non-color meanings (e.g., "*jimi*" will be said to mean "blood" before "red"). These two properties contrast with lexicalized basic color terms used in English that originate in real objects of the world (e.g., the word "orange" for the orange color). In sum, Martin's work on color categorization among the Shipibo-Konibo questions the universality of classic color categorization, and suggests that some indigenous cultures of Amazonia have developed categorization systems that strikingly contrast with ours in that they include highly context-sensitive and flexible concepts.

Many psychologists have claimed that persons are universally conceived in a dualist fashion. Martin investigated how the Shipibo-Konibo ascribe abilities ("thinking", "eating", "crying", "reasoning", etc.) to different beings (animals, plants, natural inanimate objects, artefacts); and also, how they ascribe

"ontological components" ("*shinan*" = "thinking"; "*kaya*" = "soul"; "*yora*" = "body", etc.) to the same beings. His preliminary findings contrast with the view that dualism is universal. Instead, it seems that properties of beings are ascribed in a very dynamic and context-sensitive way.

This work also speaks to recent anthropological discussions around the concept of animism. As mentioned above, some anthropologists have it that a key feature of Amazonian cultures is that they over-ascribe mental properties to animals, plants and even non-living objects: in other words, there are animists. Martin's fieldwork on components of beings paints a different picture, as he did not find the kind of hyper-mentalization described by proponents of the animistic model.

The Amazonian world is replete with "supernatural" entities, some of which are spirits and others described as masters or guardians of the animals, the river, the land, etc. (some supernatural entities are purported to be both spirits and masters). Martin studied the properties Shipibo-Konibo ascribe respectively to the master (*ibo*) of a plant (the ayahuasca vine), of a river (the Ucayali river), and of an animal (a white-lipped peccary). Interestingly, while the master of the plant is conceived as an immaterial spirit and as being the same for all the plants, the masters of the rivers and of the animals are taken to be other animals and each herd or each river possesses its own animal master. In the zoological and ecosystemic domains, the master seems to be defined by very specific ecological relationships. This work further questions some of the assumptions of approaches from both anthropology and the cognitive science of religion that over-emphasize the role of hyper-mentalization in folk intuitions about "supernatural" entities.

An influential theory developed by several authors working on the cognitive science of religion claims that the main mechanisms of supernatural thinking lie in the violation of modular and universal domain-specific knowledge. In his work, Martin pointed out some weaknesses of this theory and developed an alternative proposal. In his view, supernatural thinking is triggered by violations of randomness that cannot be explained by our ordinary causal schemas. This theory is largely inspired from previous work on randomness,

on probabilistic models of agency and on causal reasoning, and it contends that supernatural thinking is not a matter of modularity or domain-specific knowledge. This proposal challenges many influential approaches to religious thinking developed within the cognitive science of religion.

The second broad line of research that Martin was pursuing concerns about the relationship between hallucinogens, culture, and neurobiology. Psychotropic substances are pharmacological compounds that have the power of altering cognition and/or consciousness. Hallucinogens are a subclass of psychotropic compounds: they specifically alter perception by generating (visual, auditory, haptic, etc.) contents which are not veridical. Importantly, many psychostimulants (e.g., coca, khat) or sedative substances (e.g., opiates, kava) are not hallucinogenic. Martin's research was only concerned with hallucinogenic substances rather than psychotropic substances in general.

After extensively reviewing the literature on hallucinogenic compounds and their effects, Martin identified nine classes of hallucinogens: serotonergic hallucinogens (e.g., LSD, psilocybin); antimuscarinic hallucinogens (e.g., scopolamine, atropine); antihistaminergic hallucinogens (e.g., brompheniramine, diphenhydramine); κ-opioid hallucinogens (e.g., pentazocine, bremazocine); antiglutamatergic hallucinogens (e.g., ketamine, phencyclidine); cannabinoid hallucinogens (e.g., THC, nabinole); nicotinic hallucinogens (e.g., nicotine, nornicotine); anticholinesterasic hallucinogens (e.g., harmine, harmaline); and GABAergic hallucinogens (e.g., muscimol, THIP). In addition to these nine classes, he subsumed compounds whose effects tap into several neurotransmission systems, such as ibogaine and noribogaine, into a tenth class.

While most of the research investigating hallucinogens focuses on serotonergic compounds, Martin argued that critical insights on the mechanisms of drug-induced hallucinations can be gained by comparing different classes of hallucinogens. To that end, he undertook a general comparative study of the nine classes of hallucinogens, with a particular focus on the comparison between serotonergics and antimuscarinics, because these

two are the most well documented. By comparing the underlying pharmacodynamics and resulting hallucinogenic effects of serotonergics and antimuscarinics, he suggested that one-to-one mappings can be established between the neuropharmacological and the phenomenological levels.

After showing how distinct neuropharmacological mechanisms correlate with distinct phenomenological contents, Martin proposed several neurocognitive models building the bridge between the cellular and the experiential levels. For example, a key feature of antimuscarinic hallucinations is that the experience is perfectly realistic and embedded in the non-hallucinatory environment. By contrast, serotoninergic hallucinations are mostly non-realistic and extraordinary in character and they present themselves as belonging to another "plane of reality". This phenomenological contrast can be readily explained by the fact that serotonergics importantly alter the early visual cortex as well as areas underlying object construction, while antimuscarinics mainly interfere in higher areas underlying scene construction.

Martin believed that the investigation of the neuropharmacological and cognitive neuroscience of hallucinogens could also provide crucial insights for our understanding of hallucination and perception. In particular, he sought to debunk three "myths" endorsed by classical theories of hallucination and perception.

The first myth is that genuine perception can be defined as a relation between the agent and the world, whereas hallucinations are characterized by the lack of such relation. Martin's work suggests that hallucinatory states are always influenced by inputs from the world. The real difference between perception and hallucination does not lie in whether the activity of central nervous system is related to the world (influenced by the world) or not, but in how stimuli from the world are being processed. A hallucinatory state may be induced exactly by the same input (i.e., by the same relation) as the non-hallucinatory state, but what makes it hallucinatory is the way it processes that input – i.e., what kind of output it generates given a certain input.

The second myth is that there is such a thing as a perfect hallucination – i.e., a hallucinatory content that perfectly mimics a non-hallucinatory content. Martin argued that none of the different kinds of hallucination induced by his nine classes of hallucinogens ever generates perfect hallucinations. He believed this to be equally true of other types of hallucinations caused by schizophrenia, neurodegenerative diseases, dreaming, etc. While hallucinations never perfectly mimic non-hallucinatory states, they may subjectively seem to do so. How is it, then, that subjects can erroneously experience a hallucinatory state to be identical to a non-hallucinatory state while it is not? In Martin's view, this happens every time the subject lacks metacognitive insights. This metacognitive faculty, he argued, is closely related to the activation of the dorsolateral prefrontal cortex.

The third myth is that hallucinations are anomic, that is, that they are random (i.e., they can result in any kind of content and they are not governed by specific and predictable laws). This myth has been traditionally endorsed by authors such as Descartes and Malebranche. Martin's research suggests that each hallucinogen generates a specific kind of hallucinatory content and that this specificity can be elucidated by studying the cellular and neurocognitive mechanisms underpinning the hallucinatory effects. Thus, an antimuscarinic hallucinogen – unlike a serotonergic hallucinogen – will never induce extraordinary hallucinatory objects (such as a chimera). Thus, Martin argued that hallucinations are normatively constrained: they are governed by specific laws and these laws can be discovered by the neuroscience of hallucinations.

On the basis of his investigation of real hallucinations, Martin rejected both "disjunctivism" and "common factor theories" (i.e., most brands of intentionalism, sense-data, etc.) in the philosophy of perception. He rejected disjunctivism because it claims that the key difference between the perceptual state and the hallucinatory state is to be understood in terms of relation to the world (present in one case and absent in the other). This amounts to endorsing the first myth. Conversely, he rejected "common factor theories" because they endorse the second myth of "perfect hallucinations". For that matter, he also rejected predictive coding accounts of hallucinations,

which he believed to be also committed to the truth of the second myth (see the slogan "perception is a controlled hallucination").

Martin was also interested in the relationship between neuropharmacology and culture. He set out to answer the following question: when one entertains a given hallucinogenic experience, what part of the experience is caused by neuropharmacology and what part is caused by culture? From a purely biological perspective, the whole experience might be boiled down to the properties of the pharmacological compound. Conversely, from a purely culturalist perspective, the whole experience might be said to be determined by one's mindset and by the setting in which the experience is taking place.

Martin's contribution to this debate consists of highlighting two main points. First, he suggested that by comparing the role of neuropharmacology and culture in hallucinogenic experiences induced by distinct classes of compounds (serotonergics, antimuscarinics, κ-opioids, etc.), we can improve our understanding of the relative contribution made by each factor. Second, he argued that the neuropharmacology vs. culture divide is misleading because the two factors are in fact deeply intertwined. For example, he claimed that it is precisely for neuropharmacological reasons that experiences induced by serotonergic hallucinogens are so sensitive to the subject's mindset and the setting. The cultural shaping of such hallucinogenic experiences is constrained by neuropharmacology. Such cases require us to go beyond the usual neuropharmacology vs. culture divide and to develop more interactive and dynamic concepts.

A related issue is the relationship between neuropharmacology and cultural evolution. A number of authors have made strong claims regarding the role played by the consumption of hallucinogens in cultural evolution. According to them, beliefs, artifacts and institutions are crucially shaped by hallucinogenic use. In other words, hallucinogenic use results in the emergence of very specific cultural traits. To assess such claims, Martin reviewed the key cultural traits these authors take to be specifically related to hallucinogenic use. In addition, he compared traditional cultures in which hallucinogenic use was present to others in which it was absent, and

examined whether the latter lack the cultural traits purported to be found only in hallucinogenic-driven societies. This investigation revealed that the cultural traits under scrutiny are encountered in all cultures regardless of whether hallucinogenic use is present or not. Martin concluded that hallucinogens play only a limited role in cultural evolution; furthermore, their role may concern formal properties of culture, rather than its content. To take the example of beliefs, Martin thought that hallucinogens shape not so much what is believed – i.e., the content of beliefs – as the kind of attitude people entertain towards those contents. In other words, regardless of what belief is being entertained (believing in gods, in spirits, in dwarfs, etc.), hallucinogens have very specific attitudinal effects. Van Leeuwen has proposed to distinguish religious credence from factual belief, and he (with others) argued that the attitude defined as factual belief is never found among religious believers. By contrast, Martin argued that hallucinogenic use precisely enables people to switch from religious credence to factual belief and make people believe in a "factual belief mode" what they previously believed in a "religious credence mode". Hence Martin's more general claim that the role hallucinogens play in cultural evolution is more a matter of attitudinal change than a matter of content change.

In order to further investigate the interactions between hallucinogenic use and culture, Martin started developing an extensive database of hallucinogenic use across time and space, including hundreds of references from academic articles and books, explorers' diaries, and historical chronicles. Each of these references specifies which hallucinogen was used, where it was used, who was allowed to use it, what it was being used for (e.g., divinatory, political, or healing purposes), etc. This HUTHAC (Hallucinogenic Use Through History and Across Cultures) database was meant to be the first academic source exhaustively documenting hallucinogenic use through history and across cultures. The criteria used to insert data within the HUTHAC were particularly stringent. First, data included in the database concern only hallucinogenic use and not psychotropic use in general. Second, when building the HUTHAC, Martin made a distinction between use of hallucinogens at hallucinogenic doses and use of hallucinogens at sub-hallucinogenic doses (i.e., at doses too low to

induce any hallucination). Sub-hallucinogenic uses of hallucinogens are not included in the HUTHAC. Since the 1950s, outlandish philological speculations have been made about the use of hallucinogens at different epochs and in different places in the world. The HUTHAC does not include the doubtful cases of hallucinogenic use advanced by these authors. It only includes uses of hallucinogens for which there is a good amount of evidence. This evidence may be first-person observation (by an ethnographer or an explorer), third-person report (provided that the report can be considered reliable), and archeological evidence.

The main preliminary finding of Martin's research on the HUTHAC database is that hallucinogenic use is fairly rare through history and across cultures. By and large, in the last hundred years, hallucinogenic use has mainly flourished in the Americas and been absent elsewhere (with only a few exceptions). Even in the Americas, hallucinogenic use has been persistently absent from some areas – notably North America (with the exception of Southern California and the Southwest), the Southern Cone, and Eastern Brazil. Another finding is that hallucinogenic use has gradually increased through the centuries (for example, in South America, ayahuasca and mimosa use had developed only in the 18th century). These findings challenge the view that hallucinogenic use is very ancient and belongs to an ancestral shamanic tradition that would have migrated from Siberia to the New World. More broadly, they also challenge the view that most archaic cultures used hallucinogens. Instead, they support the hypothesis defended by John Cooper and Joseph Wilbert, according to which hallucinogenic use is relatively recent and related to horticultural and agricultural practices.

Some neuroscientists argue that our brain is evolutionarily wired, and as a result, it is everywhere the same and generates everywhere similar behavioral outputs. Other neuroscientists argue that the brain is wonderfully plastic and that its wiring can be extensively modified through experience. Martin believed that the study of hallucinogenic use provides us with a view of the brain that radically differs from these two stories. Regardless of whether our brain is malleable or not, what he found so fascinating about hallucinogens is that they can deeply modify behavior and experience, without modifying

anything in the underlying hardware. When someone takes LSD, for example, the wiring of the brain is exactly the same before and during the hallucinogenic experience (one can certainly observe dramatic functional changes in brain connectivity, but not any structural change). And yet, under LSD, consciousness will be so deeply altered, sounds will influence the content of vision, spatial dimensions will collapse, outlandish and folkloric characters will appear, new metaphysical beliefs will be endorsed, etc.

Martin invented the concept of *Homo neurochemicus* to highlight that most of the flexibility of human behavior stems from neurochemical changes. He believed that proponents of evolutionary psychology and of fixist models of the wiring of the brain are mistaken in that they overlook the power of neurochemistry. All the behaviors and thoughts that they take to be universal and fixed can in fact be altered by introducing the right chemical compound in the central nervous system. According to him, proponents of the plasticity of the brain are also mistaken because they think that diversity of behavior and experience can only stem from a change in the wiring of the brain; in fact, functional changes caused by neurochemical means are much more dramatic than those induced by the long-term plasticity of the brain.

It is obvious from this brief summary of Martin's work that he was a profoundly original and prolific researcher, promised to a brilliant academic career that would defy traditional disciplinary boundaries. His premature passing is all the more tragic that he cared deeply, and above all else, about answering the questions raised in these various projects.

Beyond his extensive research program, Martin was a lion-hearted colleague and loyal friend. His contribution to academia was also pedagogical. During his stay in Perú, he dedicated a lot of time and effort to mentor undergraduate students to study abroad. Some of them have now entered Master studies and Ph.D. programs in top-ranked universities worldwide. He was able to connect people across continents and ideas across disciplines.

Martin's legacy includes the formation of innovative collaborations (such as ALIUS), which will forever bear his unique signature. He was able to connect

people across continents and ideas across disciplines. Martin's legacy includes the formation of innovative collaborations (such as ALIUS), which will forever bear his unique signature. His passion for the scientific study of consciousness was inspiring, altering the trajectory of many researchers in our network and beyond. Now Martin lives on through networks: of neurons, citations, and colleagues.

Selected references

Fortier, M., Wente, A. O., Fernández Flecha, M., & Gopnik, A. (submitted). Abstract causal knowledge in Amazonia: Shipibo-Konibo horticulturalists learn multiple causality more readily than U.S. and Peruvian undergraduates.

Fortier-Davy, M., & Millière, R. (2020). The multi-dimensional approach to drug-induced states a commentary on Bayne and Carter's "dimensions of consciousness and the psychedelic state". *Neuroscience of Consciousness, 2020*(1), niaa004. https://doi.org/10.1093/nc/niaa004

Fortier, M. (2020). Enculturation without TTOM and Bayesianism without FEP: Another Bayesian theory of culture is needed. *Behavioral and Brain Sciences, 43*. https://doi.org/10.1017/S0140525X19002905

Fortier, M. (2019). Le façonnement neuropharmacologique de la culture : Anthropologie comparée des rituels à hallucinogènes sérotoninergiques et anticholinergiques. *Cahiers d'Anthropologie Sociale, 17*, 118-137. https://doi.org/10.3917/cas.017.0132

Fortier, M. (2019). L'ontologie amazonienne: animisme ou non-essentialisme? https://doi.org/10.31235/osf.io/7ntgh

Fortier, M., Kellier, D., Flecha, M. F., & Frank, M. C. (2018). Ad-hoc pragmatic implicatures among Shipibo-Konibo children in the Peruvian Amazon. https://doi.org/10.31234/osf.io/x7ad9

Fortier, M. (2018). Believing and experiencing the supernatural: Four shortcomings of the Interactive Religious Experience Model. *Religion, Brain & Behavior*, 35-38. https://doi.org/10.1080/2153599X.2018.1453531

Fortier, M. (2018). Le sens de réalité dans les expériences psychotropes : Etude comparée des hallucinogènes sérotoninergiques et anticholinergiques. *Ed. S. Baud. Histoires et usages des plantes psychotropes. Paris: Imago*, 125-184.

Fortier, M. (2018). Sense of reality, metacognition, and culture in schizophrenic and drug-induced hallucinations. In J. Proust & M. Fortier (Eds.), *Metacognitive diversity: An interdisciplinary approach*. Oxford/New York: Oxford University Press, *pp. 343–378*

Proust, J., & Fortier, M. (2018). Metacognitive diversity across cultures. In *J. Proust & M. Fortier (Eds.), Metacognitive diversity: An interdisciplinary approach.* Oxford/New York: Oxford University Press, *p. 1*

Proust, J., & Fortier, M. (Eds.). (2018). *Metacognitive diversity: An interdisciplinary approach*. Oxford University Press.

Fortier, M., & Kim, S. (2017). From the impossible to the improbable: a probabilistic account of magical beliefs and practices across development and cultures. In *The Science of Lay Theories* (pp. 265-315). Springer, Cham.

Fortier, M., & Fortier, M. (2017). Pour une science de la diversité de la conscience: Vers une redéfinition multidimensionnelle des états modifiés de conscience. *Intellectica, 67*, 27-62.

Dumas, G., Fortier, M., & González, J. (2017). Les enjeux des états modifiés de la conscience et de la cognition: limites passées et émergence de nouveaux paradigmes. *Intellectica*, 67, 27-62.

DMT in the Mammalian Brain:

A Critical Appraisal

by Charles D. Nichols & David E. Nichols

Cite as: Nichols, C. D. & Nichols, D. E. (2020). DMT in the Mammalian Brain: A Critical Appraisal by Nichols, C. D. and Nichols D. E.. *ALIUS Bulletin*, *4*, 16-22, https://doi.org/10.34700/s66k-9j57

Charles D. Nichols
cnich1@lsuhsc.edy
LSU Health Sciences Center, New Orleans, USA

David E. Nichols
drdave@purdue.edu
University of North Carolina, Chapel Hill, USA

Abstract

Recently, a publication from Dean et al. reported that N,N-dimethyl tryptamine (DMT) is synthesized in the rat brain cortex, present at levels similar to other monoamine neurotransmitters, and significantly increases in concentration at death. They further promoted the theory that DMT may serve as the causative agent for "near death experiences", which have been compared to peak psychedelic experiences. The publication certainly is interesting and suggests additional directions to explore scientifically but does not meet the bar for either claim that DMT is at functional levels in the cortex comparable to serotonin or is the "near death" neurotransmitter.

keywords: *DMT, near-death experiences, neurotransmission*

Commentary on: Dean, J. G., Liu, T., Huff, S., Sheler, B., Barker, S. A., Strassman, R. J., Wang, M. M. & Borjigin, J. (2019). Biosynthesis and extracellular concentrations of N, N-dimethyltryptamine (DMT) in Mammalian Brain. *Scientific reports*, *9*(1), 1-1 https://doi.org/10.1038/s41598-019-45812-w

N,N-dimethyl tryptamine (DMT) and the natural product admixture ayahuasca being used in shamanic rituals that contains DMT, is currently a very fashionable topic. Thousands of people worldwide have used DMT or ayahuasca in recent years under shamanic guidance or in a therapeutic setting where it appears to show benefit for antidepressant-like effects and helped with healing psychological trauma. DMT has been a trendy topic both in the scientific literature and the popular press, where several pieces have been published ascribing cure-all properties to the molecule, including proposals that it is the neurotransmitter responsible for mystical experiences in "near

death experiences" where the individual has clinically died and been resuscitated.

The recent publication by Dean et al. (2019) is the next iteration in attempts by the some of the coauthors of this publication to prove that endogenously produced DMT has important physiological functions, and in particular could be responsible for the mental effects of a near death experience. In this publication, they report that DMT is found in rat brain visual cortex, with slightly elevated levels of DMT in brains of rats that were asphyxiated. Previously, some of these authors had proposed that DMT is produced in significant amounts by the pineal gland (Barker, Borjigin et al. 2013). One of us has reviewed that hypothesis and has pointed out that the pineal gland is not capable of producing physiologically relevant amounts of DMT (Nichols 2018). The present report by Dean et al. showing that DMT could be detected in rat brain in similar amounts, regardless of whether or not the animal had a functioning pineal gland, is consistent with our earlier argument.

In a previous publication by some of these coauthors (and prior to their studies focusing on DMT), Li et al. (2015) emphasized that asphyxia generates a "brainstorm" of neurochemicals. To wit, "An immediate and sustained surge of a large set of core neurotransmitters within the cortex occurs in response to asphyxia. In both frontal and occipital cortices, a dramatic and significant surge of neurotransmitter secretion was detected for as long as 20 min of asphyxia for all neurotransmitters tested." They found that cortical levels of serotonin (5-HT) surged more than 20-fold, norepinephrine more than 30-fold, and dopamine more than 12-fold. Additionally, levels of glutamate and other neurotransmitters including acetylcholine, adenosine, aspartate, taurine, histamine, and glycine all surged within minutes after asphyxiation. Thus, it is curious that Dean et al. (2019) focuses only on DMT as an important player in brain death within this context.

Activation of brain serotonergic 5-HT2A receptors is the mechanism whereby hallucinogenic drugs such as DMT induce visual hallucinations and mystical experiences in humans (Nichols 2016). One of the metabolic pathways for the endogenous ligand of this receptor, serotonin, is N-

methylation, catalyzed by the enzyme Indoleamine N-methyltransferase (INMT), for which serotonin was its first identified substrate. Previous work by Schmid and Bohn (2010) showed that high concentrations of serotonin can induce the mouse head twitch response, an animal proxy for hallucinogenic activity, via activation of 5-HT2A receptors. Thus, the surge of serotonin alone following asphyxiation could be responsible for activation of 5-HT2A receptors and production of (the behavioral metric used in laboratory animal studies of) psychoactive effects. In addition, if INMT is responsible for production of DMT, and INMT also N-methylates serotonin (Axelrod 1962), there should be a peak in the HPLC trace for N-methylserotonin, but that is not evident in the presented HPLC traces (and these two molecules do not co-elute in HPLC analysis). Further, N-methylserotonin also activates 5-HT2A receptors to produce behavioral effects (Schmid and Bohn 2010) and this metabolite would be expected also to contribute to potential CNS effects mediated by the 5-HT2A receptor following asphyxiation. It should be mentioned that INMT is not specific for N-methylation of tryptamine, but also N-methylates a variety of other arylethylamines including, tyramine, normetanephrine, metanephrine, 3-methoxytyramine, dopamine, and octopamine (Axelrod 1962), as well as histamine (Herman, Bowsher et al. 1985). Importantly, therefore, INMT cannot be seen strictly as a proxy for DMT production.

As for the other neurotransmitters found to be dramatically induced in the cortex by asphyxia in the author's earlier work (Li, Mabrouk et al. 2015), norepinephrine has a central effect on arousal and alertness and also activates adrenergic receptors that are co-expressed on apical dendrites of cortical pyramidal cells, the same anatomic location where 5-HT2A receptors are expressed. Dopamine plays important roles in arousal, attention, cognition, and affective emotion. Increased brain glutamate concentrations can lead to out of body and hallucinogenic experiences (Gouzoulis-Mayfrank, Heekeren et al. 2005, Browne and Lucki 2013), and is a mechanism involved in the out of body experience induced by the anesthetic ketamine. With such a flood of neurochemicals, including those that can significantly impact CNS function and induce out of body experiences, it is not clear why the authors attach such importance to the relatively small increase in the amount of DMT in the

brain following asphyxiation. Apparently, it is the recurring meme that because exogenous DMT is hallucinogenic, and because it can be produced in the brain, therefore it must be important there, and have some physiological (hallucinogenic?) role.

Another explanation for out of body experiences, e.g., at near death, that the authors again fail to consider could be the production of dynorphin (DYN) and other endogenous opioid peptides. DYN and its cognate kappa-opioid receptor (KOR) play an important role in regulating stress responsiveness, motivation, and emotion (Bruchas, Land et al. 2010, Knoll and Carlezon 2010, Van't Veer and Carlezon 2013, Donahue, Landino et al. 2015). DYN 1-13 is an extremely potent kappa receptor agonist, with 0.44 nM affinity at the kappa receptor in rhesus monkey brain (Emmerson, Liu et al. 1994). Readers will appreciate that salvinorin A, the hallucinogenic component of Salvia divinorum, is a selective and extremely potent agonist at the KOR that can produce hallucinogenic and out of body experiences (Roth, Baner et al. 2002). Other endogenous opioid peptides are produced during stress and would activate other classes of opioid receptors. The authors made no attempt in their work to measure production of endorphins.

A most critical aspect lacking in the discussion was a practical understanding of receptor pharmacology. 5-HT has a 10-fold higher affinity for the 5-HT2A receptor than DMT (PDSP Ki Database). Even if we accept that DMT is present at half the levels of 5-HT in cortex under normal physiological conditions, the combination of higher 5-HT levels and higher affinity of 5-HT for the target receptor indicate that DMT will not be engaging the receptor to any appreciable degree at baseline conditions. During asphyxiation, as the authors' previous work shows, and as they interpret data here, levels of serotonin increase over 20-fold compared to only a 6-fold increase for DMT (Dean et al. Fig 4A), further widening this gap. At these comparative levels, with the 10-fold higher affinity of serotonin, 5-HT2A receptors would be saturated with serotonin and engagement of receptors by DMT in the presence of that much serotonin (and/or N-methylserotonin) would essentially be zero.

There are additional issues with experimental design and interpretation of the data presented. For example, why was microdialysis performed in the visual cortex rather than the frontal cortex, where behaviors are mediated? Why was there no validation of RNAScope results with antibodies? mRNA levels do not necessarily correlate with expressed protein levels, and to state conclusively that the enzymes INMT and AADC are co-expressed in the same cells and synthesize DMT requires validation of the presence of the enzyme proteins themselves. The statement that "DMT would be the only monoamine whose biosynthesis takes place within the cerebral cortex where it may directly influence cognitive functions of the brain" is patently false. It has been known for years that monoamine neurotransmitters, including serotonin and N-methyltryptamines, are synthesized locally within the cerebral cortex where they can influence behaviors.

Perhaps most curious, in pinealectomized animals, the peak supposedly representing 5-HT is significantly blunted in comparison (Figure 4A & B). Why is that? The pineal gland does not regulate 5-HT levels in the brain and there is no expectation that the absence of this gland would alter cortical 5-HT levels after cardiac arrest. Further, two other un-identified peaks in the HPLC trace show the most significant increases after cardiac arrest (Figure 4B). What are these peaks? The authors do not address these discrepancies, but rather only compare serotonin to DMT in pinealectomized animals.

To be very clear, we are not arguing that DMT is not produced in the cortex of the rat. Rather, even with the production of amounts of DMT indicated by the authors' data, the higher levels of serotonin, and potentially N-methylserotonin, are much more likely to induce a behavioral response through the 5-HT2A receptor. The "brainstorm" of additional neurochemicals may also be relevant to altered consciousness, and in addition, suggests the potential role of dynorphin or other endorphins, which were not measured, cannot be discounted. As Dean et al. wisely conclude, "It is unknown whether the concentrations of DMT reported in our study at cardiac arrest can elicit the effects of an exogenous psychedelic dose of DMT…" And, "the conscious states reported by NDE survivors may involve

contributions from several of the other neurotransmitters found to surge at cardiac arrest in our prior rodent study." Exactly our point here!

Science aside, a real problem with this report is that it is being taken up by the popular culture media and widely spread to a lay audience as now established dogma. Without a critical reading of the publication, advocates for the importance of endogenous DMT in the brain will and are saying, "see, we told you so." Unfortunately, it only serves to propagate a pseudoscience meme. If we take the "politics" of DMT out of the equation, and simply examine the data presented, the publication by Dean et al. certainly is interesting and suggests additional directions to explore scientifically, but does not meet the bar for either claim that DMT is at functional levels in the cortex comparable to serotonin or is the "near death" neurotransmitter.

References

Axelrod, J. (1962). The enzymatic N-methylation of serotonin and other amines. *Journal of Pharmacology and Experimental Therapeutics, 138*(1), 28-33.

Barker, S. A., Borjigin, J., Lomnicka, I., & Strassman, R. (2013). LC/MS/MS analysis of the endogenous dimethyltryptamine hallucinogens, their precursors, and major metabolites in rat pineal gland microdialysate. *Biomedical Chromatography, 27*(12), 1690-1700. https://doi.org/10.1002/bmc.2981

Browne, C. A., & Lucki, I. (2013). Antidepressant effects of ketamine: mechanisms underlying fast-acting novel antidepressants. *Frontiers in pharmacology, 4*, 161. https://doi.org/110.3389/fphar.2013.00161

Bruchas, M. R., Land, B. B., & Chavkin, C. (2010). The dynorphin/kappa opioid system as a modulator of stress-induced and pro-addictive behaviors. *Brain research, 1314*, 44-55. https://doi.org/ 10.1016/j.brainres.2009.08.062

Dean, J. G., Liu, T., Huff, S., Sheler, B., Barker, S. A., Strassman, R. J., Wang, M. M. & Borjigin, J. (2019). Biosynthesis and extracellular concentrations of N, N-dimethyltryptamine (DMT) in Mammalian Brain. *Scientific reports, 9*(1), 1-11. https://doi.org/10.1038/s41598-019-45812-w

Donahue, R. J., Landino, S. M., Golden, S. A., Carroll, F. I., Russo, S. J., & Carlezon Jr, W. A. (2015). Effects of acute and chronic social defeat stress are differentially mediated by the dynorphin/kappa-opioid receptor system. *Behavioural pharmacology, 26*(7 0 0), 654. https://doi.org/10.1097/FBP.0000000000000155

Emmerson, P. J., Liu, M. R., Woods, J. H., & Medzihradsky, F. (1994). Binding affinity and selectivity of opioids at mu, delta and kappa receptors in monkey brain membranes. *Journal of Pharmacology and Experimental Therapeutics, 271*(3), 1630-1637.

Gouzoulis-Mayfrank, E., Heekeren, K., Neukirch, A., Stoll, M., Stock, C., Obradovic, M., & Kovar, K. A. (2005). Psychological effects of (S)-ketamine and N, N-dimethyltryptamine (DMT): a double-blind, cross-over study in healthy volunteers. *Pharmacopsychiatry, 38*(06), 301-311. https://doi.org/10.1055/s-2005-916185

Herman, K. S., Bowsher, R. R., & Henry, D. P. (1985). Synthesis of N pi-methylhistamine and N alpha-methylhistamine by purified rabbit lung indolethylamine N-methyltransferase. *Journal of Biological Chemistry, 260*(22), 12336-12340.

Knoll, A. T., & Carlezon Jr, W. A. (2010). Dynorphin, stress, and depression. *Brain research, 1314*, 56-73.

Li, D., Mabrouk, O. S., Liu, T., Tian, F., Xu, G., Rengifo, S., Choi, S. J., Mathur, A., Crooks, C. P., Kennedy R. T., Wang, M. M., Ghanbari, H & Borjigin, J. (2015). Asphyxia-activated corticocardiac signaling accelerates onset of cardiac arrest. *Proceedings of the National Academy of Sciences, 112*(16), E2073-E2082. https://doi.org/10.1073/pnas.1423936112

Nichols, D. E. (2016). Psychedelics. *Pharmacological reviews, 68*(2), 264-355.

Nichols, D. E. (2018). N, N-dimethyltryptamine and the pineal gland: Separating fact from myth. *Journal of Psychopharmacology, 32*(1), 30-36. https://doi.org/10.1177/0269881117736919

Roth, B. L., Baner, K., Westkaemper, R., Siebert, D., Rice, K. C., Steinberg, S., ... & Rothman, R. B. (2002). Salvinorin A: a potent naturally occurring nonnitrogenous κ opioid selective agonist. *Proceedings of the National Academy of Sciences, 99*(18), 11934-11939. https://doi.org/10.1073/pnas.182234399

Schmid, C. L., & Bohn, L. M. (2010). Serotonin, but not N-methyltryptamines, activates the serotonin 2A receptor via a β-arrestin2/Src/Akt signaling complex in vivo. *Journal of Neuroscience, 30*(40), 13513-13524. https://doi.org/10.1523/JNEUROSCI.1665-10.2010

Van't Veer, A., & Carlezon, W. A. (2013). Role of kappa-opioid receptors in stress and anxiety-related behavior. Psychopharmacology, 229(3), 435-452. https://doi.org/10.1007/s00213-013-3195-5

Internally Generated Conscious Activity:

Reflections upon (lucid) dreaming, mind-wandering and meditation

An interview with
Benjamin Baird

by Matthieu Koroma

Cite as: Baird, B. & Koroma, M. (2020). Internally generated conscious activity: Reflections upon (lucid) dreaming, mind-wandering and meditation. An interview with Benjamin Baird by Matthieu Koroma. *ALIUS Bulletin*, *4*, 24-38, https://doi.org/10.34700/h9bs-xs28

Benjamin Baird
bbaird@wisc.edu
University of Madison-Wisconsin, USA

Matthieu Koroma
mkoroma@ens-paris-saclay.fr
Ecole Normale Supérieure, Paris, France

Abstract

Certain conscious states such as dreaming reveal that conscious activity can be to a large extent internally generated rather than being driven by sensory stimuli. In this interview, psychologist and neuroscientist Benjamin Baird discusses the developments of scientific research on these conscious phenomena including dreaming, mind-wandering and meditation and how they interrelate. Lucid dreaming, the ability to become aware that one is dreaming while in a dream, is highlighted as a unique way to gain experimental control over internally generated conscious activity during sleep.

keywords: *dreaming, lucid dreaming, mind-wandering, meditation, metacognition*

You hold a PhD in cognitive neuroscience and currently are a postdoctoral research fellow at the Wisconsin Institute for Sleep and Consciousness in the School of Medicine and Public Health at the University of Wisconsin, Madison. In your research, you have investigated topics such as mind-wandering, metacognition, meditation and dreaming. Could you introduce the scientific questions that you are targeting in your research?

Since I was young I have been interested in conscious experience and how it arises in nature. Overall, my long-term research interests extend to broad scientific questions (which I grant I will likely never see answered in my lifetime but I pursue them nonetheless!) including: how subjectivity arises in

natural systems, how the capacity for explicit self-awareness arises in the human brain, the significance of language and symbolic representation in defining human cognition, as well as issues pertaining to emergence, information, and semiotic processes in biological systems. My research to date has mostly focused on spontaneous or self-generated conscious phenomena (e.g., mind-wandering, dreaming). I believe that studies of these spontaneous states can provide unique and useful ways of approaching scientific questions regarding consciousness, including characterizing its dynamics as well as specifying the necessary conditions for its presence or absence. Broadly, I am particularly interested in self-reflective awareness, both the fundamental question of how a creature or entity can have self-reflective awareness at all, as well as more generally how we monitor our conscious states. Along these lines, I have a particular interest in lucid dreaming, and I have focused on this state in my latest research.

Dreaming and mind wandering have been proposed to rely on similar mechanisms to the extent that dreaming has been cast as "an intensified form of mind-wandering" (Fox et al., 2013). In this account, dreaming and mind wandering phenomena differ in quantitative rather than qualitative terms. Supporting this proposal, electroencephalogram (EEG) measurements have found that slow waves that are classically observed during sleep are also linked to mind-wandering experiences during wakefulness (Andrillon et al., 2019). These results blur the lines between cognitive and neural processes observed across sleep and wakefulness. Could you please explain some of the similarities and differences between dreaming during sleep and mind-wandering during wakefulness? Do you think we should study mind wandering and dreaming beyond the wakefulness vs. sleep dichotomy? What would be the advantages and pitfalls of doing so?

I think "similarities and differences" is a good way to look at it, rather than "are they the same"? They are clearly not the same. They are both sections of the brain-mind state-space if you like, so in that sense continuous, but there are important differences between dreaming and mind wandering. First, as a number of researchers have pointed out, dreams, particularly those occurring during Rapid Eye Movement (REM) sleep, often involve what might be called "full immersion", whereas mind-wandering does not. That is, in a dream you often find yourself embodied in a dream body interacting with a tridimensional (3D) multimodal virtual dream world. Even vivid daydreams

or mind-wandering episodes do not lead to this level of immersiveness, or, to borrow a term from the field of virtual reality, "presence". From this perspective, dreaming might be regarded as a full-blown world-simulation, which in this sense is more similar to our experience during the waking state as a whole, rather than just mind-wandering. This point is also related to the differences in experiential vividness of the states. For instance, in our research we have found that individuals can perform smooth pursuit eye movement tracking in REM sleep dreaming but not during visuomotor imagination during wakefulness (LaBerge, Baird & Zimbardo, 2018). These and other findings indicate that the perceptual vividness of dreaming is likely higher than that typically experienced during mind-wandering or daydreaming states.

In turn (there are many overlapping constructs here!), we have proposed that this is potentially related to the amount of disconnection or decoupling from the external environment. While we have found that mind-wandering consistently involves sensory decoupling, as revealed by event-related potentials (ERP), pupillometry, and cortical phase-locking for example, the decoupling during REM sleep is more intense, which may allow for increased vividness due to reduced competition from external stimuli. From the neuroscience perspective, an argument has been made that the states should be regarded as qualitatively the same because of overlapping neural substrates. However, while the evidence indicates that there are some brain regions that are shared between waking mind-wandering and REM sleep dreaming, there are also many differences, and overall in my view the substrate of REM sleep dreaming appears to be notably different, both in terms of neural activation patterns and neurochemistry, from that of waking mind-wandering. Finally, part of the answer to this question hinges on how we define mind-wandering. Jennifer Windt has made the important point that on a broad definition of mind-wandering as spontaneous conscious thoughts, such types of thoughts can occur in a dream! If one thing can occur within, or during, the overall state of the other, then it seems problematic to equate the two. There are many interesting similarities as well – i.e., perceptual decoupling, spontaneous internal generation, etc. – and thus overall I find the similarities and differences approach to be the most fruitful.

> “Dreams, particularly those occurring during Rapid Eye Movement (REM) sleep, often involve what might be called “full immersion”, whereas mind-wandering does not.”

Lucid dreaming is defined as a kind of dream during which dreamers are aware that they are dreaming. Stephen LaBerge provided the first experimental evidence for lucid dreaming by showing that lucid dreamers can perform I their dream an eye-movement sequence defined with the experimenter during wakefulness (LaBerge et al., 1981). You published recently a major synthesis on the advances of the neurocognitive research investigating the cerebral bases of lucid dreaming (Baird, Mota-Rolim, et al., 2019). Can you summarize the state of what we know about the neural mechanisms of lucid dreaming?

First, I would like to emphasize that research on this topic is still in its infancy and we need substantially more research before definitive conclusions can be drawn. After objectively validating lucid dreaming as a phenomenon of REM sleep using the eye signaling method, LaBerge and colleagues went on to study physiological correlates of lucid REM sleep using this technique. Their team did fantastic and rigorous research, which has unfortunately often been overlooked and not received the recognition that it deserves. One of the main findings of that initial period of research was that lucid REM sleep was associated with measures of phasic activation and autonomic nervous system arousal, including higher REM density, respiration rate and heart rate (LaBerge et al., 1986). Together these findings suggest that lucid dreams tend to occur during periods of heightened physiological activation during ongoing REM sleep.

These findings also raised the further question of whether the lucid REM sleep state was associated with a global and non-specific activation of the brain or whether it was associated with activation of specific localized brain areas or changes in specific neural oscillatory patterns. Over the past 40 years there have unfortunately only been about four EEG studies on lucid dreaming published in peer-reviewed journals. And even more unfortunately, each finds a different purported neural signature of lucid dreaming! Many of these

studies have substantial interpretive issues and limited spatial sampling of the scalp, making them hard to interpret or compare. Thus in terms of EEG all we can say right now is that the current research consists of mixed results and more research is needed. In terms of functional magnetic resonance imaging (fMRI) experiments, research is even more scant, but several recent studies have pointed to regions of the frontoparietal network as important for lucid dreaming (Dresler et al., 2012; Baird et al., 2018). Critically, however, this comes from a case report and an individual differences study and there is still no group-level fMRI. As such, that remains one of the most important goals for upcoming work. One of the most important findings we have comes from pharmacology. Specifically, we know that the probability of having a lucid dream is substantially enhanced by cholinergic stimulation during REM sleep (LaBerge, LaMarca & Baird, 2018). This fits overall with the findings noted above suggesting lucidity is associated with increased activation. In my view one of the most interesting and important next steps for research on the neurobiology of lucid dreaming therefore is to understand mechanistically why Acetylcholinesterase inhibitors have such a dramatic effect on lucid REM sleep.

Extending the approach of LaBerge, several labs around the world recently investigated the ability of experimenters to communicate with the lucid dreamer. To do so, they observed the response of the lucid dreamer to questions (e.g., arithmetic operations) using a variety of predefined "codes", such as moving the eyes in one direction or moving certain muscles of the face to say "yes" and moving the eyes in another direction or moving other muscles for "no". Which new insights can be gained by the development of an extended real-time communication with a lucid dreamer?

I think these findings are important to publish in the peer-reviewed literature as a proof of concept. Nevertheless, it is perhaps worth mentioning that many of us in the field have known for a long time that this was possible since there have been lucid dream induction devices available for decades that include the possibility for two-way interaction with the device through eye movements. I haven't yet heard a compelling argument for how this will open up new avenues of research. Most of the time lucid dreamers should be in a position to remember experimental tasks and intentions that were set in conversation with an experimenter before going to sleep, so it is hard to see

what this adds that goes beyond that in a substantially different way. Furthermore, communicating in this way with someone who is asleep and dreaming consistently runs the risk of waking him or her up. These comments should not in any way be taken to denigrate this research, as again I think it is important to publish as proof-of-concept. And perhaps there could be uses for this that I have not thought of yet. On the other hand, it does strike me that having a higher repertoire of signals on the output side could be useful for "annotating" ongoing physiological sleep recordings. For instance, my colleague Stephen LaBerge has suggested that by using something like 10 different types of eye signals a dreamer in the sleep lab could indicate that they are currently experiencing a variety of specific kinds of conscious content. By making these distinct signals, they could in real time label ongoing physiological recordings of dreams. This could be highly useful for psychophysiological studies of mind-brain relationships during sleep.

Lucid dreaming differs from dreaming in terms of metacognition, the cognitive ability of having explicit knowledge on our cognitive processes. Indeed, we know that we are dreaming when we are dreaming lucidly as contrary to regular dreaming (Kahan & LaBerge, 1994). Yet, lucid dreaming is also characterized by other mental capacities that differ from regular dreaming. For example, voluntary control is rated by the lucid dreamer as similar as wakefulness and is enhanced as compared to normal dreaming (Dresler et al., 2014). This allows for the dreamer to take the control of the course of the dreaming scenario. What can lucid dreaming teach us about the role of metacognition and more generally about the difference in cognitive activity during dreams and the waking state?

You ask an interesting and important question. In fact, the paper you cite by Kahan and LaBerge showed that non-lucid dreams are characterized by more metacognitive thought than people typically acknowledge. For instance, they found that people often experience high-level cognition in dreams, including thoughts such as "I wonder how what I just said may have caused so-and-so to feel" or "so-and-so asked me a question and I realized I didn't know the answer and I felt embarrassed", etc. These are remarkably complex thoughts that have metacognitive components i.e., directly reflecting on one's own cognition or emotional state and/or others (theory-of-mind). Nevertheless, despite this, by definition we lack a specific kind of metacognitive knowledge

in (non-lucid) dreams: namely, knowledge about the overall state of consciousness we are in. On the basis of these findings, Kahan and LaBerge argued that dreaming and wakefulness actually aren't that different in terms of ongoing thought. After all, most of us don't go around during the waking state reflecting on our state of being awake, and so it was argued that it's actually not that peculiar that we do not do this during dreaming either. However, a problem with this argument is that dreaming scenarios and content are not merely a recapitulation of our waking lives. Instead, not only do highly bizarre events occur during dreams, but I can experience myself as a totally different person, or living in a totally different century, or I could find myself on another planet! The point is that during dreams we can find ourselves in situations that are profoundly discontinuous with the rest of our waking lives and yet we still don't notice that anything is amiss. This shows how the picture is actually more complex than is typically discussed: on the one hand, we can have highly complex, including metacognitive, thoughts, and on the other hand, we can be completely disoriented, which suggests that something is profoundly different about our state of consciousness during dreams.

I think an interesting direction to pursue in resolving this apparent paradox is Tulving's notion of "autonoetic consciousness" (Tulving, 1985). In other research we find that, contrary to thoughts in the waking state and particularly mind-wandering, dreaming individuals rarely engage in episodic memory or autobiographical planning for the future. Thus one way to think about what is different during dreaming consciousness is in terms of our self-model, including our experience of ourselves as a self extended over time. Jennifer Windt and Thomas Metzinger have done excellent work in this area, and have begun to think about how many of the key changes in consciousness that occur after becoming lucid that you note above – including volitional control, episodic memory and metacognitive awareness of state – may all be related to changes in our self-model (Windt & Metzinger, 2007).

“ One way to think about what is different during dreaming consciousness is in terms of our self-model, including our experience of ourselves as a self extended over time. ”

Specifically, they have made the interesting suggestion that the specific metacognitive deficit in dreams may be directly linked to our ability to think about our self being in a current state. Similarly, as noted above, the re-instantiation of episodic memory also seems directly linked to our ability to experience selfhood, specifically the experience of ourselves over time, and volitional control is the experience of myself as an agent that can direct my attention or actions. Overall this suggests that for humans a key difference between waking and dreaming may be in how we experience ourselves as conscious subjects.

For inexperienced lucid dreamers, dream reports reveal that it is difficult to maintain the state of lucidity and attempts to control the dream scenario may lead to the loss of lucidity or to awakenings (LaBerge & Rheingold, 1990). Yet, with training, expert lucid dreamers can learn how to maintain such a state (Stumbrys et al., 2012). Such consideration can be extended to meditation as beginners might have difficulty entering the meditative state or remaining in it (Brown & Engler, 1980). How does training allow for the access and stabilization of these conscious states? How do they relate to interindividual traits and practices (Schredl & Erlacher, 2004; Baird, Riedner, et al., 2019)?

Your question points to the connections between lucid dreaming and meditation, which to my mind is an interesting topic deserving of more research. There are two distinct aspects: access and stabilization. We have found that long-term meditation practitioners report more frequent lucid dreams compared to individuals without meditation training, suggesting that there is something about meditation practice that leads to greater access to the lucid dreaming state (Baird et al., 2019). Additionally, many of the primary skills cultivated in meditation practice (particularly open-monitoring or focused-attention meditation), including stability of attention and meta-awareness, are thought to be useful in having lucid dreams. Stumbrys et al. (2015) found an intriguing association between dispositional mindfulness (a construct that broadly refers to cultivating awareness of experience in the present moment) during waking states and lucid dream frequency, but only in individuals with prior meditation experience. Although preliminary, one interpretation of this finding is that at least some types of meditation practices result in changes in trait mindfulness, or cognitive skills associated with specific aspects of mindfulness, that then

carry over into sleep and dream states, leading to increases in lucidity. Our recent research has also found a link between specific aspects of trait mindfulness and lucid dream frequency. The link between mental training and sustaining lucidity is equally interesting from a theoretical point of view, however almost no empirical work has been done to explore this connection. Dunne, Thompson and Schooler (2019) have recently argued that certain styles of mediation, for instance Tibetan Mahamudra, specifically cultivate a type of non-propositional, sustained meta-awareness that could be useful in sustaining lucidity during a dream. I would like to see this investigated in future research.

Sleep represents an interesting case for consciousness research since it consists in periods during which conscious activity is reduced, typically in NREM sleep, and a period where conscious activity qualitatively differs from wakefulness, typically in REM sleep. Yet, conscious experiences during sleep have long been considered with caution as reports about dreaming activity are available only after the sleeper has awakened and not during sleep (Malcolm, 1958; Windt, 2013). The identification of neural markers of consciousness, as well as the development of lucid dreaming research, opens novel ways to directly probe consciousness during sleep. What is the relevance of studying sleep according to you to understand consciousness?

I think there are (at least) three critical points to be made here, and then a point specifically about lucid dreaming as methodology. First, sleep is the only naturally occurring physiological state characterized by a global loss and recovery of consciousness. Thus, sleep turns out to be a useful state for building and testing theories of consciousness. Specifically, any theory of consciousness needs to both account for and be consistent with the neurophysiology of sleep states associated with consciousness. If the predictions of a neurobiological theory are inconsistent with the observation that consciousness occurs in particular brain states during sleep, then the theory needs to be revised or discarded. Having the opportunity to study a state with a different neurophysiological milieu seems a great boon to scientific research in this area. Being able to study this alternate state puts us in a stronger position to build better, more generalizable theories. Second, unlike other states, such as anesthesia-induced states of unconsciousness,

“ What lucid dreaming essentially gives us is experimental control over the dream state in a way that was previously impossible. ”

sleep affords us the unique opportunity to contrast consciousness with unconsciousness within the same overall vigilance state. There are likely many changes in the brain that take place during the shift from wakefulness to sleep or anesthesia to post-recovery wakefulness that have nothing to do with consciousness *per se*. Thus, within-state contrasts in NREM or REM sleep allow us to study specific changes in brain activity associated with the change in consciousness and to avoid confounding differences in brain activity between vigilance states that are not related to consciousness. Third, conscious states during sleep are largely decoupled from the external environment, so this fortuitously allows us to study the neural substrate of consciousness while avoiding another critical confound: namely, the neural activity associated with stimulus processing but not related to the generation of conscious percepts.

Finally, in line with the high value of sleep research for consciousness, lucid dreaming opens up a new way for us to directly study conscious content during ongoing sleep dreaming. I don’t think this fact has yet been appreciated or widely recognized. What lucid dreaming essentially gives us is experimental control over the dream state in a way that was previously impossible. As we have shown in our recent studies, lucid dreamers can perform specific tasks during dreams or invoke specific types of conscious content (e.g., LaBerge, Baird & Zimbardo, 2018). Lucid dreaming is thus an invaluable methodological tool within the cognitive neuroscience to study dreaming and by extension consciousness.

On a final note, mind-wandering has been reported for accounting about 50% of our awake life (Killingsworth & Gilbert, 2010) and sleep represents 30% of our lifetime (Iber et al., 2007). Your research topics cover thus the majority of our life. How does research on these topics change our view of mental life and the potential of the mind?

One of the things that I find most fascinating about self-generated states of consciousness is that they illustrate the profound extent to which our ongoing conscious experiences are generated internally. This is illustrated most strikingly in the case of dreams, where, as I mentioned earlier, we can experience a multimodal 3D virtual reality that can seem just as real and experientially vivid as waking life, but we know is generated completely independently of the external environment. For me personally, this has profoundly altered my view of perception. I spent most of my life as a naïve-realist, thinking that I was looking out through the windows of my eyes and seeing directly physical objects in front of me. It turns out that is not how perception works. Instead, at all times what we experience is quite literally a "virtual reality" generated by our brain. I regard this as one of the most interesting things I have ever learned. So overall, one view that emerges from studying self-generated states, in particular dreaming, is that our experience of the world is at all times endogenously generated by the ongoing dynamics of the brain, but is merely shaped by environmental input some of the time. This point has been made several times by other researchers, but I emphasize it again here because I think it does offer a profound shift in how we think about our ongoing conscious mental activities and the nature of our perceptual experience of the world.

In terms of the potentials of the mind, that is a big question. I will just mention one here. One of the ways of exploring this that has most fascinated me is examining the following question: By extending self-aware consciousness into dream and sleep states, could we tap into potentials of our minds that we haven't yet explored? This is yet another one of the reasons I have been fascinated by lucid dreaming and drawn to do scientific research on the topic. The question is, by being able to bring our "wake-like" cognitive faculties into the REM sleep dreaming state, could we explore a part of our mind-brain "state-space" that we haven't previously had access to? And could this state have unique uses, for example for things like problem-solving, art or creativity? To briefly mention an example, I met a professional composer who in his lucid dreams would find a radio and turn it on and hear symphonies being played. He would then wake up and transcribe the music that he heard in the dream into musical notation. In this way he was able to consciously use the REM sleep state for creative inspiration. I've met visual

artists that have done the same type of thing, for instance walking into an "art gallery" in their lucid dream and seeing paintings on the walls displayed, and then waking up and painting the paintings that they saw in the dream gallery. These are anecdotes and this needs to be researched. But it does raise the intriguing possibility that there may be untapped potentials of our minds that we may be able to access by extending greater awareness into the REM sleep state. Overall, I don't think we've really considered yet what the potentials could be of making the dreaming state of the brain accessible to our unique form of self-aware consciousness.

> "By extending self-aware consciousness into dream and sleep states, could we tap into potentials of our minds that we haven't yet explored?"

References

Andrillon, T., Windt, J., Silk, T., Drummond, S. P. A., Bellgrove, M. A., & Tsuchiya, N. (2019). Does the Mind Wander When the Brain Takes a Break? Local Sleep in Wakefulness, Attentional Lapses and Mind-Wandering. *Frontiers in Neuroscience*, *13*, 949. https://doi.org/10.3389/fnins.2019.00949

Baird, B., Smallwood, J., Mrazek, M. D., Kam, J. W. Y., Franklin, M. S., & Schooler, J. W. (2012). Inspired by Distraction: Mind Wandering Facilitates Creative Incubation. *Psychological Science*, *23*(10), 1117–1122. https://doi.org/10.1177/0956797612446024

Baird, B., Castelnovo, A., Gosseries, O., & Tononi, G. (2018). Frequent lucid dreaming associated with increased functional connectivity between frontopolar cortex and temporoparietal association areas. *Scientific reports*, *8*(1), 1-15. https://doi.org/ 10.1038/s41598-018-36190-w

Baird, B., Mota-Rolim, S. A., & Dresler, M. (2019). The cognitive neuroscience of lucid dreaming. *Neuroscience & Biobehavioral Reviews*, *100*, 305–323. https://doi.org/10.1016/j.neubiorev.2019.03.008

Baird, B., Riedner, B. A., Boly, M., Davidson, R. J., & Tononi, G. (2019). Increased lucid dream frequency in long-term meditators but not following

mindfulness-based stress reduction training. *Psychology of Consciousness: Theory, Research, and Practice*, *6*(1), 40–54. https://doi.org/10.1037/cns0000176

Brown, D. P., & Engler, J. (1980). The stages of mindfulness meditation: A validation study. *The Journal of Transpersonal Psychology; Palo Alto, Calif.*, *12*(2), 143–193. http://atpweb.org/jtparchive/trps-12-80-02-143.pdf

Cai, D. J., Mednick, S. A., Harrison, E. M., Kanady, J. C., & Mednick, S. C. (2009). REM, not incubation, improves creativity by priming associative networks. *Proceedings of the National Academy of Sciences*, *106*(25), 10130–10134. https://doi.org/10.1073/pnas.0900271106

Colzato, L. S., Ozturk, A., & Hommel, B. (2012). Meditate to Create: The Impact of Focused-Attention and Open-Monitoring Training on Convergent and Divergent Thinking. *Frontiers in Psychology*, *3*. https://doi.org/10.3389/fpsyg.2012.00116

Dement, W., & Kleitman, N. (1957). Cyclic variations in EEG during sleep and their relation to eye movements, body motility, and dreaming. *Electroencephalography and Clinical Neurophysiology*, *9*(4), 673–690. https://doi.org/10.1016/0013-4694(57)90088-3

Ding, X., Tang, Y.-Y., Tang, R., & Posner, M. I. (2014). Improving creativity performance by short-term meditation. *Behavioral and Brain Functions: BBF*, *10*, 9. https://doi.org/10.1186/1744-9081-10-9

Dresler, M., Wehrle, R., Spoormaker, V. I., Koch, S. P., Holsboer, F., Steiger, A., Obrig., H., Sämann, P. G., & Czisch, M. (2012). Neural correlates of dream lucidity obtained from contrasting lucid versus non-lucid REM sleep: a combined EEG/fMRI case study. *Sleep*, *35*(7), 1017-1020. https://doi.org/10.5665/sleep.1974

Dresler, M., Eibl, L., Fischer, C. F. J., Wehrle, R., Spoormaker, V. I., Steiger, A., Czisch, M., & Pawlowski, M. (2014). Volitional components of consciousness vary across wakefulness, dreaming and lucid dreaming. *Frontiers in Psychology*, *4*. https://doi.org/10.3389/fpsyg.2013.00987

Fox, K. C. R., Nijeboer, S., Solomonova, E., Domhoff, G. W., & Christoff, K. (2013). Dreaming as mind wandering: Evidence from functional neuroimaging and first-person content reports. *Frontiers in Human Neuroscience*, *7*. https://doi.org/10.3389/fnhum.2013.00412

Grecucci, A., Pappaianni, E., Siugzdaite, R., Theuninck, A., & Job, R. (2015).

Mindful Emotion Regulation: Exploring the Neurocognitive Mechanisms behind Mindfulness. *BioMed Research International*, *2015*, 1–9. https://doi.org/10.1155/2015/670724

Iber, C., Ancoli-Israel, S., Chesson, A. L., & Quan, S. F. (2007). *The AASM manual for the scoring of sleep and associated events: Rules, terminology and technical specifications (Vol. 1). Westchester, IL: American Academy of Sleep Medicine.* https://aasm.org/clinical-resources/scoring-manual/

Kahan, T. L., & LaBerge, S. (1994). Lucid dreaming as metacognition: Implications for cognitive science. *Consciousness and Cognition*, *3*, 246–264. https://doi.org/10.1111/cogs.12491

Killingsworth, M. A., & Gilbert, D. T. (2010). A Wandering Mind Is an Unhappy Mind. *Science*, *330*(6006), 932–932. https://doi.org/10.1126/science.1192439

Kircanski, K., Thompson, R. J., Sorenson, J., Sherdell, L., & Gotlib, I. H. (2018). The everyday dynamics of rumination and worry: Precipitant events and affective consequences. *Cognition and Emotion*, *32*(7), 1424–1436. https://doi.org/10.1080/02699931.2017.1278679

LaBerge, S., Baird, B., & Zimbardo, P. G. (2018). Smooth tracking of visual targets distinguishes lucid REM sleep dreaming and waking perception from imagination. *Nature communications*, *9*(1), 1-8. https://doi.org/10.1038/s41467-018-05547-0

LaBerge, S., LaMarca, K., & Baird, B. (2018). Pre-sleep treatment with galantamine stimulates lucid dreaming: A double-blind, placebo-controlled, crossover study. *PLoS One*, *13*(8), e0201246. https://doi.org/10.1371/journal.pone.0201246

La Berge, S. P., Nagel, L. E., Dement, W. C., & Zarcone, V. P. (1981). Lucid Dreaming Verified by Volitional Communication during Rem Sleep. *Perceptual and Motor Skills*, *52*(3), 727–732. https://doi.org/10.2466/pms.1981.52.3.727

LaBerge, S., & Rheingold, H. (1990). *Exploring the World of Lucid Dreaming.* New York: Ballantine Books.

Levin, R., & Nielsen, T. (2009). Nightmares, Bad Dreams, and Emotion Dysregulation: A Review and New Neurocognitive Model of Dreaming. *Current Directions in Psychological Science*, *18*(2), 84–88. https://doi.org/10.1111/j.1467-8721.2009.01614.x

Malcolm, N. (1958). *Dreaming Malcolm*.

Nielsen, T. A. (2004). Chronobiological features of dream production. *Sleep Medicine Reviews*, *8*(5), 403–424. https://doi.org/10.1016/j.smrv.2004.06.005

Schredl, M., & Erlacher, D. (2004). Lucid dreaming frequency and personality. *Personality and Individual Differences*, *37*(7), 1463–1473. https://doi.org/10.1016/j.paid.2004.02.003

Siclari, F., Baird, B., Perogamvros, L., Bernardi, G., LaRocque, J. J., Riedner, B., Boly, M., Postle, B. R., & Tononi, G. (2017). The neural correlates of dreaming. *Nature Neuroscience*, *20*(6), 872–878. https://doi.org/10.1038/nn.4545

Smallwood, J., Fishman, D. J., & Schooler, J. W. (2007). Counting the cost of an absent mind: Mind wandering as an underrecognized influence on educational performance. *Psychonomic Bulletin & Review*, *14*(2), 230–236. https://doi.org/10.3758/BF03194057

Smallwood, J., Fitzgerald, A., Miles, L. K., & Phillips, L. H. (2009). Shifting moods, wandering minds: Negative moods lead the mind to wander. *Emotion*, *9*(2), 271–276. https://doi.org/10.1037/a0014855

Tulving, E. (1985). Memory and consciousness. *Canadian Psychology/Psychologie canadienne*, *26*(1), 1. http://doi.org/10.1037/h0080017

Wagner, U., Gais, S., Haider, H., Verleger, R., & Born, J. (2004). Sleep inspires insight. *Nature*, *427*(6972), 352–355. https://doi.org/10.1038/nature02223

Walker, M. P., & van der Helm, E. (2009). Overnight therapy? The role of sleep in emotional brain processing. *Psychological Bulletin*, *135*(5), 731–748. https://doi.org/10.1037/a0016570

Windt, J. M. (2013). Reporting dream experience: Why (not) to be skeptical about dream reports. *Frontiers in Human Neuroscience*, *7*. https://doi.org/10.3389/fnhum.2013.00708

Windt, J. M., & Metzinger, T. (2007). *The philosophy of dreaming and self consciousness: what happens to the experiential subject during the dream state?*. In D. Barrett & P. McNamara (Eds.), Praeger perspectives. The new science of dreaming: Vol. 3. Cultural and theoretical perspectives (p. 193–247). Praeger Publishers/Greenwood Publishing Group. https://psycnet.apa.org/record/2007-09897-009

The Dimensions of Consciousness:

From perceptual illusions to psychedelics

An interview with
Olivia Carter

by Katrin H. Preller

Cite as: Carter, O. & Preller, K.H. (2020). The Dimensions of Consciousness: From perceptual illusions to psychedelics. An interview with Olivia Carter by Katrin H. Preller. *ALIUS Bulletin*, *4*, 40-48, https://doi.org/10.34700/mtdv-hb39

Olivia Carter
ocarter@unimelb.edu.au
University of Melbourne, Australia

Katrin H. Preller
preller@bli.uzh.ch
Psychatric Hospital University, Zürich, Switzerland

Abstract

In this interview, the psychologist and neuroscientist Olivia Carter (University of Melbourne, Australia) explains how she became interested in the field of consciousness. We discuss how her work on visual perception has led her to study the effect of psychedelics and how this has inspired a multidimensional model of consciousness. We discuss the potential contents of "higher" states of consciousness and argue that the existence of those is an unresolved question. We finish the exchange on the challenges for successful demonstration of artificial intelligence and by discussing the most important questions that the field needs to ask and answer in order to move forward.

keywords: *dimensions of consciousness, psychedelics, visual perception, artificial intelligence, content of consciousness*

You are a psychologist and neuroscientist. How did you get interested in studying consciousness? What do you think is fascinating about it and why do you think it is important?

Ever since I was a little kid, I had trouble falling asleep at night. So largely out of boredom, I remember playing games in my head trying to imagine unusual things. As I got older, I found it amazing that a brain was able to create so many different experiences. The more I learnt about neurons and the brain, the more I became fascinated by the fact that the lump of jelly in my skull could be creating all of the experiences that make me me.

For medical reasons I think it is important to understand how the brain generates a conscious experience. Either in cases of brain damage and an individual's experiences are drastically reduced or in psychiatric disorders where people can have incorrect experiences that are terrifying or confusing. I hope that the science of consciousness can progress enough to really help the lives of these different patients.

A lot of your work is focusing on visual perception. What can we learn from the visual system to understand consciousness?

It is true that much of my work has focused on vision. This has particularly involved multi-stable or ambiguous stimuli that can be validly interpreted by the brain in more than one way. These types of stimuli are very popular in consciousness science because an individual's conscious experience will fluctuate despite the stimulus remaining constant. A very striking example is binocular rivalry when two different images are simultaneously presented to the two eyes. During sustained viewing of binocular rivalry stimuli, a person will perceive one eye's image for a few seconds while the other eye's image is completely suppressed from awareness. As the images remain presented the entire time it is possible to distinguish the neural responses associated with the stimulus processing and the conscious awareness of the stimuli.

However, as I tell a lot of my students, the focus on vision is not really because vision is more interesting than other sensory modalities. It is mainly a quirk of history and technology that most of the early work in sensation focused on vision because it was possible to use prisms and physically build stimuli that could be used in vision experiments in a way that was much harder for audition and tactile experiments. So while a lot of my work has focused on vision, I share the view of many people that insights gained from vision will likely be relevant to understanding other sensory modalities. In fact, a few years ago I published a paper with some colleagues at MIT that demonstrated the first example of a perceptual illusion (perceptual rivalry) in the tactile domain by creating a tactile stimulus that mimicked the vision example (Carter, Konkle, Wang, Hayward, & Moore, 2008).

From 2002 to 2004, you worked in Zurich and conducted experiments with psilocybin. Why did you move half-way around the world to work with psychedelics?

During my PhD, I travelled to Zurich twice (each time for around 8 months) to conduct two different studies with psychedelics (Carter et al., 2004; Carter, Burr, et al., 2005; Carter, Pettigrew, et al., 2005; Carter et al., 2007; Wittmann et al., 2007). To be honest, a big motivation for the first visit was a desire to do something different and exciting and have an excuse to travel outside of Australia. By the time I completed the first study and was getting ready to head back home back to Australia, I had a better understanding of the types of research questions that could be explored with psychedelics and I could see that there was a real opportunity to do some interesting research.

What did you learn from studying psychedelics?

I still find it amazing that the majority of the effects of psychedelics are so selectively attributable to activity of the serotonin 2A receptor. It is that specificity that really interests me. As mentioned above, my main interest is in understanding how the brain generates a conscious experience. So to me, the psychedelic research really highlights the activation of this particular receptor in mediating conscious experiences. We still do not have a good understanding of why people typically hallucinate when the serotonin 2A receptor is activated, while typically they do not hallucinate when the other 200+ receptors are activated.

Psychedelics are currently being tested in the treatment of various psychiatric disorders (Bogenschutz et al., 2015; Carhart-Harris et al., 2016; Griffiths et al., 2016). However, the mechanism underlying their potential clinical efficacy is still not clear. Specifically, it is unknown whether the psychedelic experience is even necessary for the therapeutic effect. Based on your experimental and theoretical work with these substances, do you think that the alterations in consciousness induced by psychedelics are related to the therapeutic outcome?

I am not sure. Again, I think this is a very important question. Given that the drug has clear consistent effects on certain receptors and brain circuits, I definitely consider it plausible that these changes may induce therapeutic effects by themselves. I know a lot of people are suggesting that it is the nature of the experience and the quality of the associated psychotherapy sessions that are critical to the therapeutic effect rather than a direct pharmacological effect of the drug. If this turns out to be true, then it will

obviously be very important to understand which aspects of the experience and the therapy are important. It will also then be necessary to somehow control the use of these drugs to ensure that people have the desired experiences and therapy. I worry that this could be very hard to achieve if the goal is to also treat large numbers of patients with psychedelic therapy. However, I am also aware that there are many groups around the world thinking about these issues. So I will be interested to see how the clinical research unfolds.

Your recent article (Bayne & Carter, 2018) received a lot of attention. Here, you argue against unidimensional theories of consciousness. Could you summarize the idea briefly and tell us how it relates to or conflicts with other theories of consciousness?

The article was written in collaboration with the philosopher Tim Bayne who has written a number of different papers arguing against the idea that consciousness is a single thing that can go up or down along a single dimension (Bayne, 2014; Bayne, Hohwy, & Owen, 2016, 2017). In the past, people had focused on states that are typically considered "lower" like brain trauma or anesthesia. In these cases, as people move towards a state of unconsciousness, it is intuitive to imagine that everything just gets reduced somehow until the person is no-longer conscious. One problem with examples such as anesthesia or brain trauma is that they are very hard to study as the people themselves are typically unable to respond. As a result, much of the past writing has been theoretical or hypothetical. The psychedelic case allowed us to ask a different type of questions. What would it mean for an individual to be more conscious? Very quickly it becomes clear that there are multiple ways in which ways a person could be considered more conscious. An interesting conclusion of our paper was that the psychedelic state seems to lead to increases along some dimensions (such a perceptual intensity) and decreases in other dimensions (particularly those related to cognitive functions).

> "The key message is supposed to be that it is inappropriate or over-simplistic to view consciousness as a single "thing" that can be considered to exist in higher or lower states"

In your article, you therefore argue that the psychedelic state cannot be described as a "higher" state of consciousness. How was this received?

It was interesting that a number of people were angered by our paper. It is possible that people simply did not like the paper, but I think the biggest issue is that we concluded that it would be inappropriate to consider psychedelics to be a higher state of consciousness. Some people interpreted this as an argument to consider psychedelics a lower state of consciousness. That was really not our intention. The key message is supposed to be that it is inappropriate or over-simplistic to view consciousness as a single "thing" that can be considered to exist in higher or lower states. So to anyone asking the question of whether or not psychedelics lead to higher states of consciousness, I would say that this is the wrong question to ask. Psychedelics clearly lead to very unusual states of consciousness and they can be used to better understand consciousness and the underlying neural processes that determine an individual's experience from one moment to the next.

Do you think that a state in which consciousness is increased on all dimensions even exists? And if yes, how would it feel like?

That is an excellent question. Before I started working on the paper with Tim, I think I probably would have assumed there definitely would be. Now I am not so sure. The question certainly depends on what features you consider relevant to consciousness, but I think most emphasize some aspect of diversity and intensity of sensory experience while also acknowledging the importance of functional use of currently and previously experienced sensory information. For example, simple tasks such as recalling a phone number, writing a letter, or cooking a meal require us to integrate information from both our present and our past experiences. It is unclear to me that these types of tasks that require an element of concentration and focus will be enhanced or increased by increasing the diversity of sensory experience in that moment. Most people turn down music and find quiet places to work when they need to concentrate. So maybe there is an element of push and pull in terms of brain resources or functions. I think consciousness researchers would benefit from more thought given to the question of what it would mean for themselves or any other person to be *more* conscious than normal.

> I think consciousness researchers would benefit from more thought given to the question of what it would mean for themselves or any other person to be *more* conscious than normal.

On your webpage, one of the questions you are asking is "What are the factors that determine the contents of an individual's conscious experience?" What is your current answer to this question?

There are a lot of factors! I think this remains an under-explored question. I think the lessons from the paper with Tim Bayne (Bayne & Carter, 2018) are that we need to understand both the individual factors and the particular dimension of consciousness that they impact. For example, the factors that impact the experience of intense pain will be very different to the factors that enable an architect to construct a mental image of a new house design. In the case of pain, the level is likely to depend on stimulation of the specific receptors and neural pathways involved. Pain is an extremely important experience, but it is relatively simple and reflects events happening in the moment. In the case of complex building designs an individual will likely need to draw on years of educational training and experience with a mix of some creativity and consideration of budgetary factors and other practical constraints. If we want to truly understand the brain processes involved in conscious experiences, then we need to fully understand how different factors contribute to very different types of human experiences.

In another recent article (Carter et al., 2018) you ask a different but somewhat related question: "What would constitute successful demonstration of artificial intelligence?" What is your answer to this question?

I still do not have a good answer to that question unfortunately. Again, I think the community needs to think more about the different capacities that we would expect to see and consider which capacities are already demonstrated in existing systems. I think an equally important question is which capacities we care more about. It might be the case that we have greater concerns about capacities that are not as complex. For example, concepts like free-will are often considered uniquely human, but there are

already artificial intelligence (AI) systems that can formulate their own goals. In the last few years, there has been a lot more discussion around these questions. My guess is that these conversations will advance a lot in the next few years.

Speaking of the future, in which direction does the field need to move to make progress in understanding consciousness?

I think we need to start focusing on the details. Consciousness is a great umbrella term. I do not think we need a new definition of consciousness, but rather we need to be clearer about which aspects or dimensions of consciousness that individual experiments are testing.

Do you have a recommendation for young researchers who want to study consciousness?

My advice relates a lot to my answer to the previous question. I think the term consciousness covers a lot of things. It is worth thinking about what aspects of consciousness you are really interested in. There is a lot of research at the moment looking at different brain processes and their relation to consciousness. There are mathematicians and computer scientists trying to understand and simulate the complexity and information integration achieved by the human brain. At the other extreme there are philosophers really trying to understand the experiential side of consciousness. I think all of these areas of research are very important, but they are also very different and require different types of skills and I think they trigger different types of curiosities. So I think if someone believes they would like to study consciousness, I would suggest they start by reflecting more on what aspect of consciousness they would like to better understand.

“ I do not think we need a new definition of consciousness, but rather we need to be clearer about which aspects of dimensions of consciousness that individual experiments are testing. ”

References

Bayne, T. (2014). The Unity of Consciousness. *Analysis*, *74*(3), 488-490. https://doi.org/10.1093/analys/anu054

Bayne, T., & Carter, O. (2018). Dimensions of consciousness and the psychedelic state. *Neurosci Conscious*, *2018*(1), niy008. https://doi.org/10.1093/nc/niy008

Bayne, T., Hohwy, J., & Owen, A. M. (2016). Are There Levels of Consciousness? *Trends in Cognitive Sciences*, *20*(6), 405-413. https://doi.org/10.1016/j.tics.2016.03.009

Bayne, T., Hohwy, J., & Owen, A. M. (2017). Reforming the taxonomy in disorders of consciousness. *Annals of Neurology*, *82*(6), 866-872. https://doi.org/10.1002/ana.25088

Bogenschutz, M. P., Forcehimes, A. A., Pommy, J. A., Wilcox, C. E., Barbosa, P. C., & Strassman, R. J. (2015). Psilocybin-assisted treatment for alcohol dependence: a proof-of-concept study. *J Psychopharmacol*, *29*(3), 289-299. https://doi.org/10.1177/0269881114565144

Carhart-Harris, R. L., Bolstridge, M., Rucker, J., Day, C. M., Erritzoe, D., Kaelen, M., Bloomfield, M., Rickard, J. A., Forbes, B., Feilding, A., Taylor, D., Piling, S., Curran, V. H., Nutt, D. J. (2016). Psilocybin with psychological support for treatment-resistant depression: an open-label feasibility study. *Lancet Psychiatry*, *3*(7), 619-627. https://doi.org/10.1016/S2215-0366(16)30065-7

Carter, O., Hohwy, J., van Boxtel, J., Lamme, V., Block, N., Koch, C., & Tsuchiya, N. (2018). Conscious machines: Defining questions. *Science*, *359*(6374), 400-400. https://doi.org/10.1126/science.aar4163

Carter, O., Konkle, T., Wang, Q., Hayward, V., & Moore, C. (2008). Tactile rivalry demonstrated with an ambiguous apparent-motion quartet. *Curr Biol*, *18*(14), 1050-1054. https://doi.org/10.1016/j.cub.2008.06.027

Carter, O. L., Burr, D. C., Pettigrew, J. D., Wallis, G. M., Hasler, F., & Vollenweider, F. X. (2005). Using psilocybin to investigate the relationship between attention, working memory, and the serotonin 1A and 2A receptors. *Journal of Cognitive Neuroscience*, *17*(10), 1497-1508. https://doi.org/10.1162/089892905774597191

Carter, O. L., Hasler, F., Pettigrew, J. D., Wallis, G. M., Liu, G. B., & Vollenweider, F. X. (2007). Psilocybin links binocular rivalry switch rate to attention and

subjective arousal levels in humans. *Psychopharmacology (Berl), 195*(3), 415-424. https://doi.org/10.1007/s00213-007-0930-9

Carter, O. L., Pettigrew, J. D., Burr, D. C., Alais, D., Hasler, F., & Vollenweider, F. X. (2004). Psilocybin impairs high-level but not low-level motion perception. *Neuroreport, 15*(12), 1947-1951. https://doi.org/10.1097/00001756-200408260-00023

Carter, O. L., Pettigrew, J. D., Hasler, F., Wallis, G. M., Liu, G. B., Hell, D., & Vollenweider, F. X. (2005). Modulating the rate and rhythmicity of perceptual rivalry alternations with the mixed 5-HT2A and 5-HT1A agonist psilocybin. *Neuropsychopharmacology, 30*(6), 1154-1162. https://doi.org/10.1038/sj.npp.1300621

Griffiths, R. R., Johnson, M. W., Carducci, M. A., Umbricht, A., Richards, W. A., Richards, B. D., . . . Klinedinst, M. A. (2016). Psilocybin produces substantial and sustained decreases in depression and anxiety in patients with life-threatening cancer: A randomized double-blind trial. *Journal of Psychopharmacology, 30*(12), 1181-1197. https://doi.org/10.1177/0269881116675513

Wittmann, M., Carter, O., Hasler, F., Cahn, B. R., Grimberg, U., Spring, P., Hell, D., Flohr, H., Vollenweider, F. X. (2007). Effects of psilocybin on time perception and temporal control of behaviour in humans. *J Psychopharmacol, 21*(1), 50-64. https://doi.org/10.1177/0269881106065859

The Use of Psychedelics in the Treatment of Disorders of Consciousness

An interview with Olivia Gosseries

by Charlotte Martial

Cite as: Gosseries, O. & Martial, C. (2020). The use of psychedelics in the treatment of disorders of consciousness. An interview of Olivia Gosseries by Charlotte Martial. *ALIUS Bulletin*, *4*, 50-66, https://doi.org/10.34700/gv0w-y052

Olivia Gosseries
ogosseries@uliege.be
Coma Science Group,
GIGA Consciousness,
University of Liege, Belgium

Charlotte Martial
cmartial@uliege.be
Coma Science Group,
GIGA Consciousness,
University of Liege, Belgium

Abstract

In this interview, we discuss the use of psychedelic drugs as a promising treatment in disorders of consciousness. Psilocybin, a classic psychedelic, is currently undergoing substantial clinical investigations in healthy volunteers, but also in clinical populations. Recently, experts in the field of psychedelics have addressed the attractive possibility to use such psychedelics on patients suffering from disorders of consciousness. Building on her empirical and theoretical research on disorders of consciousness, Olivia Gosseries gives us her opinion. Implementing rigorous clinical trials with psychedelics on patients with disorders of consciousness will allow their clinical efficacy to be tested. We finish the interview by briefly addressing the ethical and legal challenges and discussing other related non-pathological modified states of consciousness.

keywords: *psychedelic*, *disorder of consciousness*, *treatment*, *psilocybin*, *patient*

You are well known for your impressive work on disorders of consciousness (DOC) and you have recently become the co-director of the Coma Science Group (with Dr. Aurore Thibaut), succeeding Prof. Steven Laureys who founded the Coma Science Group. Can you please introduce yourself and explain what brought you to work with patients with DOC?

Yes of course. I am a neuropsychologist who pursued a PhD in biomedical and pharmaceutical sciences. Currently I am studying altered states of consciousness, with a focus on diagnosis, prognosis and treatments of brain-injured patients with DOC.

In 2005, I spent 6 months at the University of Quebec in Montreal as part of the Socrates II program (Erasmus). During this time, I participated in a neuropsychology workshop for which I went to a rehabilitation institute every Wednesday. The aim was to interact with patients and provide a diagnosis, without having access to their medical files. My first patient did not talk and did not respond to me. I was very confused. How can I give a diagnosis in such a situation? The diagnosis of this patient was mixed aphasia: he did not understand me (sensory aphasia) and he was not able to speak (motor aphasia). Meeting with many patients over the weeks, I noticed that they all had previously fallen into a coma. Some patients had vague memories of the coma, some reported dream-like experiences and some had no recollection. I was very intrigued by this peculiar state: what is it like to be in a coma? Where is the mind in such a state? Why do some patients recover while some don't?

When I came back to Belgium, I wanted to do my internship and master's thesis on this topic. I contacted Dr. Steven Laureys who openly welcomed me and I started working as an undergrad at the Coma Science Group in 2006. Seeing patients in a coma in the intensive care units was a unique and shattering experience. I was really compelled to find a cure for coma. After completing my master's, I spent 4 months at the Moss Rehabilitation Institute in Philadelphia with Dr. John Whyte to get more familiar with the daily recovery of patients with DOC. I learned that most of these patients (if not all) do not remember their time in rehabilitation, and that a good proportion of them do not recover well. Time post-injury, etiology and the patients' clinical status upon admission were however predictors of early recovery (Whyte et al., 2009). After this enriched experience, I was ready to pursue a PhD with comatose patients. Following the PhD and several postdoctoral positions, I recently became a research associate at F.R.S-FNRS and co-director of the Coma Science Group.

After 14 years working in the field, there is still no cure for coma. My aim is to find one (wishful thinking) and to continue improving the care of patients with DOC. I also want to contribute to the understanding of human consciousness and promote public awareness of this fascinating topic. For a review on how to measure consciousness in DOC, see Gosseries et al. (2014a),

and our video Dance my PhD 2018 - the (un)conscious brain (https://www.youtube.com/watch?v=eYMmVNei2Hc&t).

Recently, you published a review in 'The Lancet Neurology' (Thibaut et al., 2019) highlighting the rarity of effective treatment options for patients with DOC. What are the main reasons that efficient treatments are rare in these situations? In addition, have we observed an increase of interests in the search for treatments over the past few years?

Yes indeed, we crucially lack effective therapeutic options for patients with DOC. I think there are four main reasons at play: 1) DOC is considered a rather new and rare disease, 2) it is a challenging patient population, 3) previous work mostly focused on diagnosis and prognosis rather than treatment specifically, and 4) we need to understand the mechanisms of consciousness recovery before developing targeted treatments.

Before the 1950's, patients with severe brain injuries would die. With the advent of the ventilators, such patients can now live, and the definition of death changed from cardiac death (now called "clinical death") to brain death (official definition of death). Unfortunately, a proportion of these patients may remain alive but without recovering. DOC is an umbrella term that includes coma (no arousal, no awareness), unresponsive wakefulness syndrome (previously known as vegetative state; wakefulness but reflexive behaviors only), and minimally conscious state (wakefulness and behavioral signs of consciousness without functional communication) (Bodart et al., 2013). Coma is an acute state in which patients will never open their eyes and it lasts more than one hour, up to a few weeks. In contrast, patients with an unresponsive wakefulness syndrome or in a minimally conscious state have their eyes open, and these states can be acute, prolonged (more than a month), or chronic (more than several months, sometimes decades).

Chronic DOC is considered a rare disease because it affects between 0.2 and 17 individuals out of 100 000 in Europe and the United-States. DOC are not classified as complications of common diseases but as pathologies per se, because they have specific International Classification of Diseases (ICD) codes. Patients with DOC are sadly often neglected by healthcare systems and private companies have little financial incentive to develop new

> “We consequently have an urgent need to investigate therapies with the support of universities and other organizations”

treatments for such rare patients. We consequently have an urgent need to investigate therapies with the support of universities and other organizations.

The second reason for the scarcity of treatment options is that management of patients with DOC is very challenging because they cannot communicate, have severe motor disability and are dependent on others for all care. Most current rehabilitation therapies need active participation of the patients, which is not possible with these patients.

The third reason is that previous work has mainly investigated diagnostic and prognostic indicators of consciousness. In medicine, before prescribing a treatment, we need to know what the diagnosis is, and in our case an accurate diagnosis of the level of consciousness (“is the person conscious?”). This can be a hard task because we primarily use behavior to infer consciousness, but responsiveness does not always equate consciousness. This may lead to a high rate of misdiagnosis (Stender et al., 2014), which can in turn lead to inadequate medical decisions, such as withdrawal of life-sustaining care. The second difficulty that arises after the diagnosis is the prognosis (“is the patient going to recover?”). With time, the patient may recover spontaneously, however we only think of treatments when patients do not recover. After more than three decades of research, we are slowly moving towards treatments.

The last and probably most important reason why we have not yet found a cure is that to treat a condition, we need to understand it. Only when we will comprehend the mechanisms of consciousness recovery, will we be able to develop effective and specific therapies. So far all the tested treatments are repurposed, meaning that they are available on the market for other pathologies and we try them with our patients. Here is an illustration based on our recent review (Thibaut et al., 2019) to show you what has been tested so far with pharmacological and brain stimulation treatments,

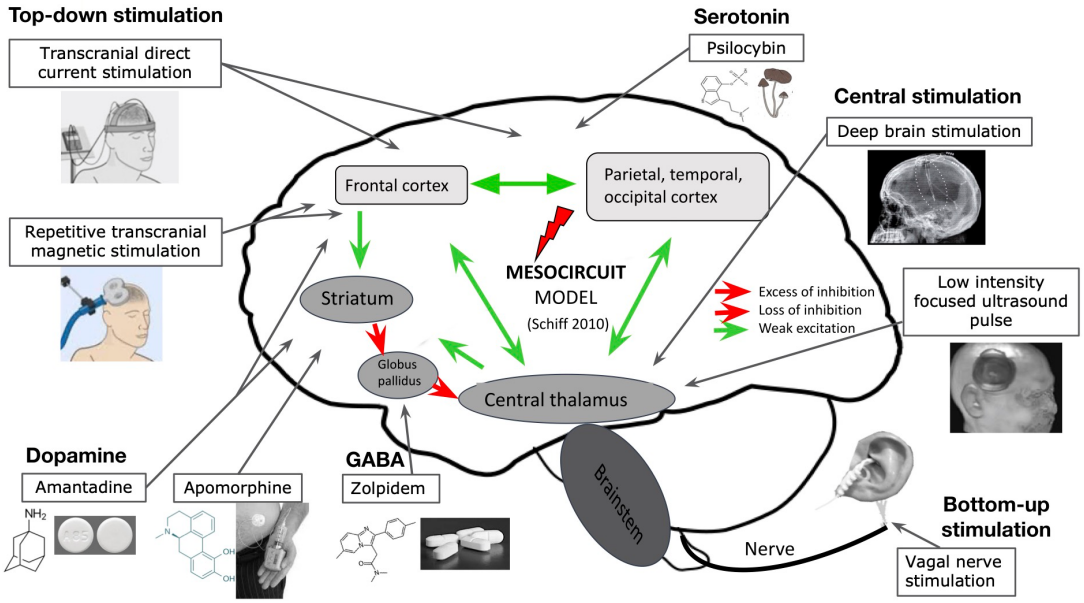

Figure 1. Pharmacological and brain stimulation treatments in patients with DOC.

and how we think they act on the damaged brain based on the mesocircuit model (Schiff, 2010). I took the liberty to add the psilocybin as this is the main topic of this interview.

To answer your second question, it is clear that recently there is more interest in treatments. There are many new research groups working on coma and related states, and many new clinical trials are being conducted around the world. The interest for DOC treatment was probably pushed forward with the publication of a landmark paper on amantadine in 2012 (Giacino et al., 2012). Very recently, the Curing Coma Campaign has been launched by the Neurocritical Care Society with the aim to develop and implement coma treatment strategies. This is the first global public health initiative created to tackle the unifying concept of coma as a treatable medical entity. I have great hope of success and invite everyone interested to join this campaign (www.curingcoma.org).

> “ Only when we will comprehend the mechanisms of consciousness recovery, will we be able to develop effective and specific therapies ”

Of note, most published treatments studies on DOC are open-label and case reports, which means that results need to be interpreted with caution and cannot be translated into clinical practice. To evaluate the effectiveness of a treatment, we need randomized controlled trials with robust designs on large samples to take into account biases, such as spontaneous recovery. Only a handful of these trials have been published so far but many are on their way (36 are currently registered as pending clinical trials).

In 2019, Scott and Carhart-Harris published a paper discussing the potential capacity of classical psychedelic, psilocybin, to increase consciousness in patients with DOC (Scott & Carhart-Harris, 2019). What was your first thought when learning about this suggestion?

This was a great moment! I was very excited and shared the paper with my team. I have been thinking about this possibility in the past but without acting on it. Discussing the paper with you and other colleagues made it clear that we had to give it a try, and that we should collaborate with the authors (which you ended up doing). The use of psilocybin to treat disorders of consciousness is an innovative, auspicious and original idea of treatment but it is certainly challenging ethically and legally.

At the Coma Science Group, we have conducted several studies on anesthesia, including administering ketamine to participants who ended up completely unresponsive at the bedside but yet reported psychedelic experiences afterwards (Sarasso et al., 2015). One of my favorite techniques to investigate brain activity is transcranial magnetic stimulation combined with electroencephalography (TMS-EEG) (for review, see Gosseries et al., 2014b). We showed in collaboration with the team of Dr M. Massimini at the University of Milan, that under ketamine, the brain reacts to the stimulation in a complex and widespread manner, like it does in normal wakefulness. In comparison, during propofol or xenon sedation, the brain reacts in a slow,

“ The use of psilocybin to treat disorders of consciousness is an innovative, auspicious and original idea of treatment but it is certainly challenging ethically and legally. ”

stereotypical and non-complex way and subjects do not recall any subjective experience afterwards. Similarly, unresponsive patients show a local and slow responses to TMS while patients in minimally conscious states show a differentiated, complex and broad responses (Rosanova et al., 2012). Going a step further, we developed the perturbational complexity index (PCI) that uses the normalized Lempel-Ziv complexity to compress the spatio-temporal pattern of cortical activation information into one number, with a threshold for consciousness above 0.31 (Casali et al., 2013, Casarotto et al., 2016). In ketamine, the PCI is high, as in healthy wakefulness, in REM sleep and in patients in minimally conscious states. In propofol and xenon sedation, PCI is low, as in non-rapid eye movement sleep (with no dream reports upon awakening) and in unresponsive patients. These findings highlight that the loss of consciousness is linked to drops in complexity in brain activity. In previous works, we also showed decreases in neural complexity in DOC patients using EEG entropy (Gosseries et al., 2011, Piarulli et al., 2016).

On the other hand, studies by other groups showed that psilocybin, a serotonin 2A-receptor agonist (5-HT2A/C), increases brain complexity in healthy subjects (Schartner et al., 2017, Varley et al., 2020). So it is a fair hypothesis that psilocybin may be a plausible awakening drug for DOC patients that would restore the loss of brain complexity and ultimately improve the patients' state and responsiveness.

Since then, have you heard some of your colleagues from the DOC field discussing that option?

Internally, yes we discussed this option and developed a study protocol that we plan to start in 2021, hopefully in collaboration with experts in the field of psychedelics. I have however not heard much interest in this drug in other groups working with DOC. This may be because of practical and legal reasons, as many institutional review boards would probably not allow the use of illegal drugs in a fragile population. There is also a panoply of other legal drugs (e.g., pitolisant, d-cycloserine) that could be tested to potentially increase patients' responsiveness. Implementing clinical trials with these drugs may be easier and more feasible at this time.

This is nevertheless an excellent opportunity to bridge the fields of DOC and psychedelics. As of now, these are separated research topics despite both being considered as non-ordinary states of consciousness. In this regard, the Mind & Life Europe (https://www.mindandlife-europe.org/) organized a workshop with research groups from different fields to start comparing states of consciousness such as meditation, hypnosis, psychedelics and trance from a first- and third-person perspective. This was a very interesting event.

What do you think about it now? Do you think the neurocognitive mechanisms sustaining the action of psilocybin could increase consciousness awareness in those patients?

This is a complicated question because we know little about the mechanisms of action of psilocybin and even less about the mechanisms of induced recovery of consciousness. It is thus difficult to predict what will be the response at the group level. I would hypothesize that psilocybin will improve the level of consciousness in a majority of patients with DOC along with increased brain complexity. I expect diffuse changes in the brain, especially in the frontal areas (where there are many serotonin receptors), in the default mode network and in the dorsal-attention network, based on previous works (Beliveau et al., 2017, Carhart-Harris and Friston, 2019, Varley et al., 2020). Psilocybin may modulate activity in the cortico-striato-thalamocortical loop, which would reactivate the consciousness network. If no clinical improvement is observed after psilocybin intake, a heightened brain activity may still be noticed. This could reflect an increase in internal awareness that is not possible to assess externally, as this is the case during ketamine sedation. But if patients recover afterwards, we could ask if they remember the experience, and if so, it would suggest disconnected consciousness at the time. Patients have however severe memory impairments so we cannot conclude anything in the absence of such reports. If we observe an increase in brain activity without improvement of responsiveness, a second explanation could be that brain complexity may be independent of consciousness, which would call into question the relationship between consciousness and brain complexity (see also Pal et al., 2019), a key element in consciousness model research (Koch et al., 2016).

> So it is a fair hypothesis that psilocybin may be a plausible awakening drug for DOC patients that would restore the loss of brain complexity and ultimately improve the patients' state and responsiveness.

In a recent dream, I gave psilocybin to a woman with DOC, and after severe side effects (I thought she was having a heart attack), she started speaking, which she had not done since her accident. So I do dream psilocybin can become a successful new awakening drug, but if not, as a Cartesian scientist, I would favor the explanation of disconnected consciousness over the brain-mind separation.

I would like to take this opportunity to relate the case of zolpidem, to show how hard it can be to predict drug effects. Zolpidem, also known as Stilnox or Ambien, is a non-benzodiazepine hypnotic drug that acts as an agonist of the inhibitory GABA receptor and binds to the GABA-A receptor chloride channel. It is commonly prescribed as a sleep inducer. In 2000, a case of zolpidem-induced awareness was reported for the first time in a patient, who after being diagnosed in an unresponsive wakefulness syndrome for 3 years, suddenly 'awoke' and began speaking 15 minutes after receiving the drug (Clauss et al., 2000). Zolpidem was initially administered to the patient to help him sleep because he was restless at night. Since then, temporary improvements of arousal, awareness and cognitive abilities have been shown in other patients with DOC, sometimes with drastic changes such as the ability to eat and walk, but just for the time of the medication (around 4 hours). This behavioral awakening is accompanied by increased metabolism in frontal regions and decreases in EEG power and coherence in low frequencies (6-10Hz) (Williams et al., 2013, Chatelle et al., 2014). Unfortunately, only 5% of patients with DOC show such paradoxical responses (Whyte et al., 2014), and it is unknown why them and not others. Increasing our knowledge of physiological processes will help to spur drug discovery, but as in the case of zolpidem, serendipity occasionally leads to important drug discoveries. This phenomenon has also been observed in the discovery of penicillin and played a role in the development of many psychotropic drugs which later helped to shape the field of psychiatry.

While there is currently no effective treatment, there are some promising approaches, particularly for patients in a minimally conscious state (MCS). Do you think psilocybin could be more efficient for patients in a MCS, compared to patients with an unresponsive wakefulness syndrome (UWS)? If so, why?

Amantadine is the only treatment that was recommended by the American practice guidelines in 2018 for patients with DOC between 4 and 16 weeks after a traumatic brain injury (Giacino et al., 2018). Zolpidem and brain stimulation techniques may also be good candidates for some patients. For instance, a previous study by Thibaut and colleagues (2014) showed that half of the patients in minimally conscious state responded to transcranial direct current stimulation targeting the prefrontal cortex, whereas only a few unresponsive patients showed new signs of consciousness. This suggests that patients may need to be above some particular threshold of conscious awareness, with a minimum of brain complexity to benefit from brain stimulation. Note that improvements in these cases remain moderate, with responders showing new behaviors such as response to command, visual pursuit or localization to pain. Some patients may benefit from therapeutic interventions even years after the brain injury (Estraneo et al., 2010).

Another promising approach that has not yet been tested is to combine different treatments (e.g., amantadine with repetitive TMS). Model-driven treatments should be developed using dynamical whole-brain computational models to understand the fundamental mechanisms of consciousness recovery. Finding biomarkers to predict responsiveness will ultimately help personalized treatment based on patients' individual profile. As of now, for brain stimulation, we know that one should stimulate on (partially) preserved structural brain area to induce brain responses and behavioral outputs (Gosseries et al., 2015, Thibaut et al., 2015).

Considering all this, one could speculate that psilocybin would be more beneficial for patients in minimally conscious states, as they have more preserved brains than unresponsive patients. On the other side, psilocybin has been shown to increase brain complexity, so it may work in unresponsive patients who specifically lack complex activity. Psilocybin might reintroduce complexity in their brain and thus responsiveness. It is also possible that a minimum of brain complexity is needed, and patients who are diagnosed at

the bedside as unresponsive but who show brain activity compatible with the minimally conscious state would be the ones who respond the most to the drug. This category of patients is referred to as being in a non-behavioral minimally conscious state* (aka MCS*) (Gosseries et al., 2014c) or experiencing a cognitive motor dissociation (Schiff, 2015). Only time will tell what will be the results of psilocybin in DOC patients.

According to you, what are the two most important challenges inherent in the testing of psychedelics in patients with DOC?

The two main challenges are related to ethical and legal issues. Working with DOC patients is already an ethical challenge as they cannot communicate and thus cannot give their informed consent. We consequently rely on their legal guardians to make decisions for them. If patients recover functional communication during psychedelic treatment, they will be able to share their experience, give their consent (or not), and even possibly make their own decisions. One concern with psychedelics is that they can produce psychotropic aversive effects, which may be frightening for the patients. We do not want to induce a negative psychological impact with this intervention, and we have to evaluate the risk-benefit ratio. We have to develop novel treatments that are the most beneficent and the least harmful. The beneficent obligation that professional caregivers have in these situations calls for the creation of a care ethos that reflects the principle of respect of persons for these patients (Blain-Moraes et al., 2018).

To avoid "bad trips" with potential harmful effects, paying attention to contextual factors will be crucial to create a relaxing setting (e.g., using decoration in the room, a comforting glow). Giving clear information to the patients and their families, and having them at the bedside during the testing, will also increase the likelihood of a safe and positive experience. Another concern is what do we do next if the patient recovers during treatment? The effects will most probably be temporary and the patients will have to take the drug repeatedly (like with zolpidem, some patients take it three times a day to eat). But psychedelics are currently illegal. It is going to be a long battle before the routine implementation of such drugs in the clinic. The legal aspect of studying psychedelic drugs in DOC patients will also need to be considered and approved by institutional review boards.

At the University Hospital of Liège, you are in frequent contact with the families of DOC patients. Do you think they might be willing to test this potential treatment on their relatives?

Yes, most families want to try everything possible for their loved ones. They might however have concerns about side effects, so we have to evaluate the risk-benefit ratio before proposing treatment. Families have to be fully informed before providing their consent. They should act in the best interest of the patient, and think of what (s)he would have liked. If there were (official or informal) advance directives, families should respect the wishes of the patients. Collaborating with families in such context is paramount and they should participate in the study and provide feedback.

Next to your interest in DOC, you have recently published an article about cognitive trance. Do you think there is a link between cognitive trance and psychedelics in terms of phenomenology and/or underlying neurophysiological mechanisms?

Great question and the short answer is yes, definitively. With my colleague Dr. Audrey Vanhaudenhuyse and others, we started investigating the effects of trance at the phenomenological and neurophysiological level over a year ago. There are certainly commonalities between trance and psychedelics. Some trance experts report anecdotally that trance is like psychedelic without the ingestion of drugs. Here is an excerpt of a trance experience related by Corine Sombrun who was trained in Mongolia and who came to our laboratory: "I saw a little ant and then I was this ant. I climbed in a tree and I fell from it. After, I had visions of insects and big lizards. I experienced a transformation again, with the feeling of becoming something else, like an iguana. Then my tongue started to come out with the sensation of a turtle's tongue. After, there were the hisses of snakes, I went through all the reptiles. I had a feeling of joy, I wanted to laugh. (...) Then it was pure joy, total happiness and a huge expansion of my perception of self" (Gosseries et al., 2019). Regarding the neurophysiological mechanisms, we have conducted an EEG study on a group of trance experts and the analyses are underway. As in psychedelics, we expect to measure elevated brain activity and complexity, along with an augmented consciousness phenomenology.

Another link to make is between trance, psychedelic and near-death experience (NDE), as some features of NDE seem to be common, such as the feeling of extrasensory perception and ineffability (Martial et al., under revision; for a recent review see Martial et al., 2020).

The ultimate nature of consciousness is still full of mysteries but it is evident that deepening our knowledge of all the possible states of consciousness can only increase our understanding of the human mind and brain. I think trance studies will open a new window of research possibilities, and maybe one day we will be able to use it to cure coma.

> “The ultimate nature of consciousness is still full of mysteries but it is evident that deepening our knowledge of all the possible states of consciousness can only increase our understanding of the human mind and brain. I think trance studies will open a new window of research possibilities, and maybe one day we will be able to use it to cure coma.”

References

Beliveau, V., Ganz, M., Feng, L., Ozenne, B., Højgaard, L., Fisher, P. M., ... & Knudsen, G. M. (2017). A high-resolution in vivo atlas of the human brain's serotonin system. *Journal of Neuroscience, 37(1)*, 120-128. https://doi.org/10.1523/JNEUROSCI.2830-16.2016

Blain-Moraes, S., Racine, E., & Mashour, G. A. (2018). Consciousness and Personhood in Medical Care. *Frontiers in human neuroscience, 12*, 306. https://doi.org/10.3389/fnhum.2018.00306

Bodart, O., Laureys, S., & Gosseries, O. (2013). Coma and disorders of consciousness: scientific advances and practical considerations for clinicians. In *Seminars in neurology* (Vol. 33, pp. 83-90). Thieme Medical Publishing Inc.. https://doi.org/10.1055/s-0033-1348965

Carhart-Harris, R. L., & Friston, K. J. (2019). REBUS and the anarchic brain: toward a unified model of the brain action of psychedelics. *Pharmacological*

reviews, 71(3), 316-344. https://doi.org/10.1124/pr.118.017160

Casali, A. G. *, Gosseries, O. *, Rosanova, M., Boly, M., Sarasso, S., Casali, K. R., ... & Massimini, M. (2013). A theoretically based index of consciousness independent of sensory processing and behavior. *Science translational medicine*, 5(198), 198ra105-198ra105. https://doi.org/10.1126/scitranslmed.3006294

Casarotto, S., Comanducci, A., Rosanova, M., Sarasso, S., Fecchio, M., Napolitani, M., ... & Gosseries, O. (2016). Stratification of unresponsive patients by an independently validated index of brain complexity. *Annals of neurology, 80(5),* 718-729. https://doi.org/10.1002/ana.24779

Chatelle, C. *, Thibaut, A. *, Gosseries, O., Bruno, M. A., Demertzi, A., Bernard, C., ... & Laureys, S. (2014). Changes in cerebral metabolism in patients with a minimally conscious state responding to zolpidem. *Frontiers in Human Neuroscience, 8,* 917. https://doi.org/10.3389/fnhum.2014.00917

Clauss, R. P., Güldenpfennig, W. M., Nel, H. W., Sathekge, M. M., & Venkannagari, R. R. (2000). Extraordinary arousal from semi-comatose state on zolpidem. *South African Medical Journal, 90(1),* 68-72. https://www.ajol.info/index.php/samj/article/download/140537/130277

Estraneo, A., Moretta, P., Loreto, V., Lanzillo, B., Santoro, L., & Trojano, L. (2010). Late recovery after traumatic, anoxic, or hemorrhagic long-lasting vegetative state. *Neurology, 75(3),* 239-245. https://doi.org/10.1212/WNL.0b013e3181e8e8cc

Giacino, J. T., Katz, D. I., Schiff, N. D., Whyte, J., Ashman, E. J., Ashwal, S., ... & Nakase-Richardson, R. (2018). Practice guideline update recommendations summary: disorders of consciousness: report of the Guideline Development, Dissemination, and Implementation Subcommittee of the American Academy of Neurology; the American Congress of Rehabilitation Medicine; and the National Institute on Disability, Independent Living, and Rehabilitation Research. *Archives of physical medicine and rehabilitation, 99(9),* 1699-1709. https://doi.org/10.1016/j.apmr.2018.07.001

Giacino, J. T. *, Whyte, J. *, Bagiella, E., Kalmar, K., Childs, N., Khademi, A., ... & Yablon, S. A. (2012). Placebo-controlled trial of amantadine for severe traumatic brain injury. *New England Journal of Medicine, 366(9),* 819-826. https://doi.org/10.1212/WNL.0b013e3181e8e8cc

Gosseries, O., Di, H., Laureys, S., & Boly, M. (2014). Measuring consciousness in

severely damaged brains. *Annual Review of Neuroscience*, *37*, 457-478. https://doi.org/10.1146/annurev-neuro-062012-170339

Gosseries, O. *, Fecchio, M. *, Wolff, A., Sanz, L., Sombrun, C., Vanhaudenhuyse, A., & Laureys, S. (2019). Behavioural and brain responses in cognitive trance: A TMS-EEG case study. *Clinical neurophysiology*, *131(2)*, 586-588. https://doi.org/10.1016/j.clinph.2019.11.011

Gosseries, O., Sarasso, S., Casarotto, S., Boly, M., Schnakers, C., Napolitani, M., ... & Laureys, S. (2015). On the cerebral origin of EEG responses to TMS: insights from severe cortical lesions. *Brain stimulation*, *8(1)*, 142-149. https://doi.org/10.1016/j.brs.2014.10.008

Gosseries, O. *, Schnakers, C. *, Ledoux, D., Vanhaudenhuyse, A., Bruno, M. A., Demertzi, A., ... & Peeters, E. (2011). Automated EEG entropy measurements in coma, vegetative state/unresponsive wakefulness syndrome and minimally conscious state. *Functional neurology*, *26(1)*, 25.

Gosseries, O., Thibaut, A., Boly, M., Rosanova, M., Massimini, M., & Laureys, S. (2014, February). Assessing consciousness in coma and related states using transcranial magnetic stimulation combined with electroencephalography. In Annales françaises d'anesthésie et de réanimation (Vol. 33, No. 2, pp. 65-71). Elsevier Masson. http://doi.org/10.1016/j.annfar.2013.11.002

Gosseries, O., Zasler, N. D., & Laureys, S. (2014). Recent advances in disorders of consciousness: focus on the diagnosis. *Brain injury*, *28(9)*, 1141-1150. https://doi.org/10.3109/02699052.2014.920522

Koch, C., Massimini, M., Boly, M., & Tononi, G. (2016). Neural correlates of consciousness: progress and problems. *Nature Reviews Neuroscience*, *17(5)*, 307-321. https://doi.org/10.1038/nrn.2016.22

Martial, C., Cassol, H., Laureys, S., & Gosseries, O. (2020). Near-death experience as a probe to explore (disconnected) consciousness. *Trends in Cognitive Sciences*, *24(3)*, 173-183. https://doi.org/10.1016/j.tics.2019.12.010

Martial, C., J. Simon, N. Puttaert, O. Gosseries, V. Charland-Verville, B. Greyson, ... H. Cassol (Under revision). The Near-Death Experience Content (NDE-C) scale: development and psychometric validation.

Pal, D., Li, D., Dean, J. G., Brito, M. A., Liu, T., Fryzel, A. M., ... & Mashour, G. A. (2020). Level of consciousness is dissociable from electroencephalographic measures of cortical connectivity, slow oscillations, and complexity. *Journal*

of Neuroscience, 40(3), 605-618. https://doi.org/10.1523/JNEUROSCI.1910-19.2019

Piarulli, A., Bergamasco, M., Thibaut, A., Cologan, V., Gosseries, O., & Laureys, S. (2016). EEG ultradian rhythmicity differences in disorders of consciousness during wakefulness. *Journal of neurology, 263(9),* 1746-1760. https://doi.org/10.1007/s00415-016-8196-y

Rosanova, M. *, Gosseries, O. *, Casarotto, S., Boly, M., Casali, A. G., Bruno, M. A., ... & Massimini, M. (2012). Recovery of cortical effective connectivity and recovery of consciousness in vegetative patients. *Brain, 135(4),* 1308-1320. https://doi.org/10.1093/brain/awr340

Sarasso, S. *, Boly, M. *, Napolitani, M., Gosseries, O., Charland-Verville, V., Casarotto, S., ... & Rex, S. (2015). Consciousness and complexity during unresponsiveness induced by propofol, xenon, and ketamine. *Current Biology, 25(23),* 3099-3105. https://doi.org/10.1016/j.cub.2015.10.014

Schartner, M. M., Carhart-Harris, R. L., Barrett, A. B., Seth, A. K., & Muthukumaraswamy, S. D. (2017). Increased spontaneous MEG signal diversity for psychoactive doses of ketamine, LSD and psilocybin. *Scientific reports, 7,* 46421. https://doi.org/10.1038/srep46421

Schiff, N. D. (2010). Recovery of consciousness after brain injury: a mesocircuit hypothesis. *Trends in neurosciences, 33(1),* 1-9. https://doi.org/10.1016/j.tins.2009.11.002

Schiff, N. D. (2015). Cognitive motor dissociation following severe brain injuries. *JAMA neurology, 72(12),* 1413-1415. https://doi.org/10.1001/jamaneurol.2015.2899

Scott, G., & Carhart-Harris, R. L. (2019). Psychedelics as a treatment for disorders of consciousness. *Neuroscience of consciousness, 2019(1),* niz003. https://doi.org/10.1093/nc/niz003

Stender, J. *, Gosseries, O. *, Bruno, M. A., Charland-Verville, V., Vanhaudenhuyse, A., Demertzi, A., ... & Soddu, A. (2014). Diagnostic precision of PET imaging and functional MRI in disorders of consciousness: a clinical validation study. *The Lancet, 384(9942),* 514-522. https://doi.org/10.1016/S0140-6736(14)60042-8

Thibaut, A., Bruno, M. A., Ledoux, D., Demertzi, A., & Laureys, S. (2014). tDCS in patients with disorders of consciousness: sham-controlled randomized

double-blind study. *Neurology, 82(13),* 1112-1118. https://doi.org/10.1212/WNL.0000000000000260

Thibaut, A. *, Di Perri, C. *, Chatelle, C., Bruno, M. A., Bahri, M. A., Wannez, S., ... & Hustinx, R. (2015). Clinical response to tDCS depends on residual brain metabolism and grey matter integrity in patients with minimally conscious state. *Brain Stimulation, 8(6),* 1116-1123. https://doi.org/10.1016/j.brs.2015.07.024

Thibaut, A., Schiff, N., Giacino, J., Laureys, S., & Gosseries, O. (2019). Therapeutic interventions in patients with prolonged disorders of consciousness. *The Lancet Neurology, 18(6),* 600-614. https://doi.org/10.1016/S1474-4422(19)30031-6

Varley, T. F., Carhart-Harris, R., Roseman, L., Menon, D. K., & Stamatakis, E. A. (2020). Serotonergic psychedelics LSD & psilocybin increase the fractal dimension of cortical brain activity in spatial and temporal domains. *Neuroimage, 220,* 117049. https://doi.org/10.1016/j.neuroimage.2020.117049

Whyte, J. *, Gosseries, O. *, Chervoneva, I., DiPasquale, M. C., Giacino, J., Kalmar, K., ... & Mercer, W. (2009). Predictors of short-term outcome in brain-injured patients with disorders of consciousness. *Progress in brain research, 177,* 63-72. https://doi.org/10.1016/S0079-6123(09)17706-3

Whyte, J., Rajan, R., Rosenbaum, A., Katz, D., Kalmar, K., Seel, R., ... & Kaelin, D. (2014). Zolpidem and restoration of consciousness. *American Journal of Physical Medicine & Rehabilitation, 93(2),* 101-113. https://doi.org/10.1097/PHM.0000000000000069

Williams, S. T., Conte, M. M., Goldfine, A. M., Noirhomme, Q., Gosseries, O., Thonnard, M., ... & Laureys, S. (2013). Common resting brain dynamics indicate a possible mechanism underlying zolpidem response in severe brain injury. *Elife, 2,* e01157. https://doi.org/10.7554/eLife.01157

* Contributed equally

The Neuroscience of Decision Making

An interview with Timothy Hanks

by Alexandra Mikhailova & Daniel A. Friedman

Cite as: Hanks, T., Mikhailova, A., & Friedman, D.A. (2020). The Neuroscience of Decision Making. An interview with Timothy Hanks by Alexandra Mikhailova and Daniel A. Friedman. *ALIUS Bulletin*, *4*, 68-80, https://doi.org/10.34700/8pg4-0h12

Timothy Hanks
thanks@ucdavis.edu
Center for Neuroscience,
UC Davis, CA, USA

Alexandra Mikhailova
amikhailova@ucdavis.edu
Center for Neuroscience,
UC Davis, CA, USA

Daniel A. Friedman
dafriedman@ucdavis.edu
Department of Entomology
& Nematology, UC Davis, CA, USA

Abstract

In this interview, Professor Tim Hanks discusses topics related to neuroscience, decision making, philosophy, and science as a career. Hanks explores how ideas from computational neuroscience have helped him set his own research agenda and also navigate everyday situations. The way the brain makes decisions is deeply intertwined with topics such as free will, conscious awareness, and mental health. In order to productively study such diverse topics related to decision making, Hanks recommends an integrative approach that draws on multiple types of experiments and model systems, with an eye towards clinical deployment. His approach builds on various scientific frameworks, while also reminding us to stay open-minded about what the future of neuroscience may look like or bring.

keywords: *decision making, Bayesian, free will, mental health, attention*

To quote your book chapter from 'Neurobiology of Decision Making': "Thus looking for the right questions is just another kind of decision" (Shadlen et al., 2008). This quote resonates with folk sayings such as "the answer is in the question", and also with computational perspectives on decision making that highlight the role of (beliefs about) prior beliefs in controlling how evidential stimuli play a role in decision making. As a scientist, how did you come to study decision making? What decision making rules or heuristics were helpful for you during your development as a professional and researcher?

I was always drawn to the questions that I found most mysterious, with those of mental experience among the top of the list. These are the questions that I find most engaging, that keep me thinking at night. What

> “ The most helpful decision-making advice I have applied in my life and career is to consider every choice in a Bayesian sense ”

provided the final push in the direction of neuroscience was the belief that these questions were ripe for finding answers. For decision making in particular, I was drawn in by a fascination with the specific research being conducted in that area that I learned about as an undergraduate student. That's what led me to do pre- and post-doctoral research in the labs of Mike Shadlen and Carlos Brody. I was enthusiastic enough about both of their research programs that I can still recall in vivid detail the joy of positive email correspondence with both before joining their labs. So, I followed the path of my strongest interests. The most important advice I can give is to follow a path that results in sustained day-to-day internal satisfaction, and to take efforts to be honest with yourself about what that involves.

Along my path, I have also carefully considered what would be most beneficial in the long term to allow me to continue to follow my interests. That's a lot harder to determine. I think the most helpful decision-making advice I have applied in my life and career is to consider every choice in a Bayesian sense by trying to use evidence to estimate probabilities or at the very least, to explicitly consider uncertainty. This can obviously help to avoid overconfidence, but it can also help to overcome paralyzing doubt. Most importantly, it gives a principled foundation for choosing courses of action. And that's ultimately why we have a brain, to shape our actions.

Every day we are consciously aware of deliberative decisions that we make (e.g. which clothes to wear, what to write in an email to a colleague), rising like islands of awareness from a vast ocean of subconscious decisions (e.g. physiological decisions related to blood pressure, heart rate, oculomotor tracking, postural balance). There are also situations where subconsciously-decided preferences can apparently be steered by conscious input, for example in the case of a deliberative pacific response to an aggressive stimulus. How do you think about the role of conscious awareness in decision making across these domains, and especially in the case of decisions where phenomenologically we seem to “have a say” in the output?

What neural events might distinguish the awareness of a decision from ones made subconsciously? Is conscious awareness necessary, sufficient, or differentially associated with decision making tasks (Ganupuru et al., 2019) or meta-cognitive assessments? How do we coherently generate research agendas that try to bridge the experience of conscious awareness, with neural or molecular measurements?

Let me start with the "have a say" component of the question because that is where I think we have the best opportunity for scientific leverage. This relates to the question of free will, which I think is often posed in a way that stifles progress on finding answers. If the question is, "Do we have free will?", the answer depends on how we define free will. I think a better starting point is to ask the question, "To what extent do we have free will?" This promotes inquiry of what it is that we actually have. As a scientist studying decision making, I can see a clear roadmap towards answering this by describing the neural mechanisms that underlie our decisions. It's my belief that when that description is complete, it will provide a satisfying and compelling answer to the question of free will much in the way that our current understanding of biology, including genetics and evolutionary theory, provide a satisfying and compelling answer to the question of life.

To make this more concrete, let me give some examples of what I think these types of explanations could look like. Through the work of many researchers across our field, we have found that commitment to a decision in many situations can be described as occurring through the accumulation of evidence to a threshold level or bound (Gold & Shadlen, 2007; Hanks & Summerfield, 2017). The evidence is represented in neural activity and the bound has been hypothesized to be applied to the level of neural activity. A higher bound would require more evidence for decision commitment and a lower bound would require less evidence. In other words, a person could respond differently to the same evidence depending on the level of the

> "Perhaps in following this roadmap, we will find aspects of decision making that defy explanation based on our current scientific understanding, or that are better explained through new scientific conceptual frameworks"

decision bound. The bound is one mechanism for exerting will. Understanding the factors that go into how a person sets their decision bound then tells us something about how that person's will can be set. This is just one example. There are many other neural mechanisms for exerting will in decision making. I think that when we have discovered and understood those mechanisms, that will provide a compelling account of the extent to which we have free will.

In reading this, one might think I am describing an approach aimed at yielding a deterministic account that will be reducible to our current understanding of physical laws, but I don't think this is guaranteed. Perhaps in following this roadmap, we will find aspects of decision making that defy explanation based on our current scientific understanding, or that are better explained through new scientific conceptual frameworks.

This is where I suspect consciousness will come to the fore. Currently, we have no satisfying and compelling scientific explanation of consciousness. Yet, as your question suggests, our internal experience gives us the impression that consciousness matters for our decisions. If that impression is correct, then in trying to fully explain the mechanisms of our decisions, I suspect that eventually we will hit that wall. The hope is that how we hit the wall will reveal how our current scientific conceptions need to be refined to better understand consciousness.

As we pursue this scientific path, I think it is important to keep in mind that consciousness is not a singular type. In this respect, I find it useful to consider the evolutionary history of consciousness. Are different types of consciousness (e.g., visual, auditory, volitional) the product of divergent or convergent evolution? It seems possible that different forms of consciousness derived from convergent evolution with distinct types of relationships with neural or molecular measurements. Even if they derived from something more akin to divergent evolution from a common proto-consciousness, I would still not be surprised if there were distinct types of relationships with neural or molecular measurements for different forms of consciousness. That could prove very helpful in our quest for deeper understanding because the commonalities between these relationships would potentially reveal principles of a more general nature.

> “As we pursue this scientific path, I think it is important to keep in mind that consciousness is not a singular type”

Your work on the speed vs. accuracy trade-off in the brains of macaques was testing the hypothesis that fast decisions require a lower threshold of evidence (Hanks et al., 2014). Your findings suggested that control of speed vs. accuracy may be exerted through changes in decision related neural activity itself, rather than through changes in the threshold applied to such neural activity to terminate a decision. This is consistent with Bayesian frameworks of decision making in the brain - i.e. that variation in internal neural states or priors can weigh on the outcome of decisions, given the same task stimuli. How do you think about the statistical or computational perspectives on decision making, in regard to decision processes that can be measured in the lab, as well as in the context of the complex long-term decisions that humans have to make?

Yes, that is very nicely put, and connects directly to the point I was making above. We studied the neural mechanisms that govern the tradeoff between the speed and accuracy of decision making because we hope that they provide more general insight into flexible control of decision commitment. In theory, control over the speed vs. accuracy trade-off could occur in a variety of ways. A lower “bound” for decision commitment might be implemented directly through less neural activity being needed to trigger a choice. Alternatively, it might be implemented with the same level of neural activity needed to trigger a choice, but additional internally-generated drive to this neural activity. In the both cases, there is less external drive needed, and therefore decisions are made more hastily. We found the mechanism to be more internal drive, what is sometimes referred to as an “urgency” signal to respond separate from the evidence.

One thing that I really like about this result is that it seems to be a fairly general mechanism for cognitive control of the brain. For decision making, it has been replicated in multiple other studies from different labs and different decision tasks (Heitz et al., 2014; Thura & Cisek, 2016). There are other situations where neuroscientists have described boosted neural activity

> “Even if you only care about the computational type of understanding or only care about the implementation type of understanding, studying both will probably help you more effectively understand either alone”

associated with cognitive control, perhaps the most prominent being with selective attention (Squire et al., 2013). In that case, attention to a particular part or feature of one's environment can boost activity of neurons selective for that part or feature. While we don't know exactly how this activity is boosted for attention or decision making, it's intriguing to consider that they may be related mechanisms. Studying these mechanisms will help us know more about what can be controlled through cognition and how. For example, let's say that after it is started, the boosting can't be suppressed for some period of time due to a limitation of the mechanism. In that case, we could truly say that a person does not have the free will to reduce their decision bound and increase it again within that time limiting span. I pose that merely as an example. I do not expect things to be so clear cut. Instead, we will probably find greater flexibility in cognitive control than this example, but it may be the case that some forms of it are easier to achieve than others.

To address the final part of your question, the Bayesian framework provides a theoretical foundation for understanding decision making, and we relate much of our work to it (Beck et al., 2008). This “computational” perspective is often contrasted with the brain's “implementation” within neural circuits. When considering this distinction, it is important to realize that the boundary may not always be clear cut and that knowledge of either can be helpful for understanding both. In other words, even if you only care about the computational type of understanding or only care about the implementation type of understanding, studying both will probably help you more effectively understand either alone.

Your lab developed a new change detection task based on a stream of auditory clicks generated by a stochastic Poisson process (Johnson et al., 2017), which allows you to measure the temporal weighting of sensory evidence when subjects employ different decision-making rules (thus

finessing the trade-off between false-positives and false-negatives). The article concluded that "changes in decision stopping rules did not alter the temporal weighting of sensory evidence on the decision in a systematic way. Instead, it altered the magnitude of evidence needed to trigger a choice." Unlike a forced-choice task, where participants have a window of time to make a selection, your task allows subjects to "decide when to decide". What other applications of this task would like to explore? What do you think is gained or lost when we isolate decision making outcomes (e.g. bets) from the time required to make such decisions? How do you think the methodological tension between fixed-time and free-time decision making tasks could be retrospectively or prospectively integrated?

We began to study change detection tasks because it allowed us to address an important set of questions neglected by much of the previous work in the field, including my own previous work. For many of the most common tasks used to study the neural mechanisms of decision making, evidence should be weighted equally across time. One of the major questions has been to understand how that consistent weighting of new evidence can be achieved, and tremendous progress has been made with answering that (Brody & Hanks, 2016). However, there are many situations where consistent weighting of evidence across time is detrimental. In any situation with change or instability, recent evidence is more informative than evidence from the more distant past. So, there can actually be an advantage to "forgetting" – or at least, not relying as much on information gathered further in the past (Glaze et al., 2015; Radillo et al., 2017; Piet et al., 2018).

In trying to take the next step to understand how our brain weighs evidence across time, we have been systematically exploring and characterizing our capacity to evaluate evidence on different timescales. Clearly, people are capable of changing the timescale of evidence evaluation for decision making. But what are the limits of this flexibility? This is what we are trying to determine. We have been using tasks that push this flexibility to see if people can simultaneously evaluate the same sources of evidence across different timescales. If you are making a quick decision about one aspect of something, does it limit your ability to combine information simultaneously over longer periods of time to make a more careful decision about some other aspect of the same thing?

These questions seem most important in situations where we must determine when to decide, but I would argue that many "forced-choice" decision-making tasks satisfy similar conditions. When you are "forced" to decide at a fixed cue, this can be thought of as driving a very strong urgency signal to respond, similar to what I described previously. In this light, perhaps we should not have been surprised to find the mechanism for changing the speed vs. accuracy tradeoff that we discussed above involved boosting neural responses rather than reducing the level of neural activity needed for commitment, because no such reduction is seen for cued responses in forced-choice tasks. Likewise, when a cue to respond comes relatively late, decision commitment may have occurred already (Kiani et al., 2008). One of the advantages of the cognitive nature of decision making is that it does not have to be tied directly to the outside world, what one of my former advisors (Mike Shadlen) would often describe as "freedom from immediacy".

Although it would stand to reason that the dynamics of the environment dictate the optimal timescale (Ossmy et al., 2013; Glaze et al., 2015; Radillo et al., 2017; Piet et al., 2018), you find that the brain represents and utilizes multiple timescales of evidence evaluation during deliberation (Ganpuru et al., 2019). What implications does this finding have for cellular mechanisms of decision making? Does this imply a brain region involved in the task can have heterogenous evidence weighting, or that multiple brain regions are responsible for heterogeneous timescales?

We believe this speaks to the architecture of neural circuits that support decision making. It establishes minimum capacities for information processing that any neural mechanism of evidence evaluation must support. Many existing models of decision processing have relied on precisely tuned circuits that can effectively evaluate evidence over a set timescale (Wong & Wang, 2006). Changing the timescales can often be achieved by altering the tuning of these circuits, but that is not enough to support multiple timescales simultaneously. Intriguingly, alternative models already exist that can support this. One such class of models uses a processing cascade with progressively longer timescales of evaluation for later nodes in the network (Goldman, 2009; Scott et al., 2017). Under this scheme, different timescales are represented by different network nodes without requiring any alterations in tuning, so it could naturally support simultaneous deliberation across

multiple timescales. We have not yet shown this to be the case, but we think it is a good hypothesis. In theory, these nodes could be distributed across different brain regions or be distinct subcomponents within brain regions related to decision making, and this is also a question we are pursuing.

Disease states can influence evidence accumulation, resulting in too much or too little exploration in various domains. How is decision making affected in schizophrenia and depression? Does this point to underlying mechanistic etiological similarities among disease states, or computational attractor states that can be induced by a variety of causes? What are promising cellular mechanisms or affected circuits to investigate in relationship to decision-making deficits?

Impaired decision making is observed in almost all mental disorders, including schizophrenia and depression, as you mention. Furthermore, these impairments can often severely impact quality of life, and some of the largest negative impacts from many mental disorders come from poor decision making. I saw this firsthand with my dad, who suffered from an atypical variant of Alzheimer's disease that presented as frontal-temporal dementia. The biggest impacts on his life early on in the disease progression were with decision making.

While decision making can be impaired in a variety of ways with different disorders, there do seem to be common modes of impairment that can occur across disorders. What we are trying to do now is exactly in line with what you propose of applying our computational models to describe these modes of impairment and link them to underlying neural mechanisms. We have begun to collaborate with UC Davis Conte Center led by Drs. Cameron Carter and Kimberly McAllister to apply this to schizophrenia patients and animal models aimed at figuring out neural mechanisms involved. We believe that corticostriatal circuits may play a central role in the impairments to decision making that come about with schizophrenia. With Dr. Randy O'Reilly, we have developed a computational model for impaired decision making in schizophrenia that makes specific predictions for neural mechanisms impacted in corticostriatal circuits, and we are beginning to test those predictions.

> “Attentional allocation can be viewed as a decision process, and a reduced bound that would explain impulsivity would also explain lower thresholds for changing attentional focus”

We have also been collaborating with Dr. Johannes Hell to study neural mechanisms of attentional control that might be affected in ADHD. One of the most interesting aspects of ADHD is that it results in impairments to not only attention, but it also leads to more impulsivity. What is the connection between the two? Through our work on decision commitment, we have a good handle on mechanisms that may be involved for impulsivity, and again, corticostriatal circuits are implicated. I believe that similar circuit mechanisms may also underlie attentional deficits. In particular, individuals with ADHD are often able to reap the benefits of attention when they are attending to something, but the problem is that they are more easily distracted. It isn't so much a problem with attentional enhancement of perception, but rather one of allocating those attentional resources. Attentional allocation can be viewed as a decision process, and a reduced bound that would explain impulsivity would also explain lower thresholds for changing attentional focus – the distractibility that affects those with ADHD. We think that corticostriatal circuit mechanisms may explain both. With Dr. Hell, we are trying to determine specific cellular mechanisms that may be involved.

What is the minimal number of neurons required to 'make a decision'? More generally, do neurons as a cell type participate in a special type of decision making relative to non-neuronal cells or perhaps even non-cells? In other words, how should we study the "decision making" processes of bacteria, computers, neural and non-neural cell types?

I take an inclusive view on what it means to "make a decision", but with the caveat that there are important differences among the class of processes considered as such. Under this view, a single neuron is certainly capable of making a decision, but without the same range of complexity as a large network of neurons. Likewise, while neurons and neural networks participate in a variety of types of decision making, they do not do so in a necessarily

unique way. Non-neural cell types, computers, etc. may also participate in similar types of decision making, but this needs to be considered on a case-by-case basis. For instance, we can describe accumulating information to a threshold over a flexible timescale from milliseconds to minutes as a specific type of decision making. Human brains can do this, but it is not known what minimal architecture of brain tissue is needed to accomplish it. Computers can easily be programmed to do it, but it becomes more challenging as we expand the domain that we are considering to more complex types of decision making. It would also be interesting to know the capacity of other systems like bacteria. Intriguingly, for this example, there are suggestions that non-neuronal cells could play a role. In particular, the glial cells (non-neuronal cells of the nervous system that actually outnumber neurons) known as astrocytes have been shown to accumulate information over 10s of seconds for decision making in zebrafish, with some indications of flexibility in that timescale. In that circuit, the astrocytes seem to work in concert with neurons by accumulating evidence from neural input over a longer timescale and then influencing other neurons.

I believe one of the keys for studying decision making in any setting is striving for a clear description of what type of decision process one is studying. This is one of the more important benefits of using a computational approach because it gives a framework for rigorous description. This has a parallel to the discussions above about free will. Rather than asking, "Is [X] making a decision?", I think it is better to ask, "In what way is [X] making a decision?" Reflexive movements may still be considered decisions in some sense, but they are very different types of decisions than those involving reflective deliberation. They invoke different forms of computation and different complexity of neural circuit processing.

References

Beck J. M., Ma W. J., Kiani R., Hanks T., Churchland A. K., Roitman J., Shadlen, M. N., & Pouget, A. (2008) Probabilistic population codes for Bayesian decision making. *Neuron*, *60*. 1142–1152. https://doi.org/10.1016/j.neuron.2008.09.021

Brody C. D., Hanks T. D. (2016). Neural underpinnings of the evidence accumulator. *Current Opinion Neurobiology*. *37*. 149–157. https://doi.org/10.1016/j.conb.2016.01.003

Ganupuru P., Goldring A. B., Harun R., Hanks T. D. (2019). Flexibility of Timescales of Evidence Evaluation for Decision Making. *Current Biology*. *29*: 2091–2097.e4. https://doi.org/10.1016/j.cub.2019.05.037

Glaze C. M., Kable J. W., Gold J. I. (2015). Normative evidence accumulation in unpredictable environments. *eLife*. *4*. https://doi.org/10.7554/eLife.08825

Gold J. I., Shadlen M. N. (2007). The neural basis of decision making. *Annual Review Neuroscience*. *30*. 535–574. https://doi.org/10.1146/annurev.neuro.29.051605.113038

Goldman M. S. (2009). Memory without feedback in a neural network. *Neuron*. *61*. 621–634. https://doi.org/10.1016/j.neuron.2008.12.012

Hanks T., Kiani R., Shadlen M. N. (2014). A neural mechanism of speed-accuracy tradeoff in macaque area LIP. *eLife*. *3*. https://doi.org/10.7554/eLife.02260

Hanks T. D., Summerfield C. (2017). Perceptual Decision Making in Rodents, Monkeys, and Humans. *Neuron*. *93*. 15–31. https://doi.org/10.1016/j.neuron.2016.12.003

Heitz R. P., Schall J. D. (2012). Neural mechanisms of speed-accuracy tradeoff. *Neuron*. *76*. 616–628. https://doi.org/10.1016/j.neuron.2012.08.030

Johnson B., Verma R., Sun M., Hanks T. D. (2017). Characterization of decision commitment rule alterations during an auditory change detection task. *Journal of Neurophysiology*. *118*. 2526–2536. https://doi.org/10.1152/jn.00071.2017

Kiani R., Hanks T. D., Shadlen M. N. (2008). Bounded integration in parietal cortex underlies decisions even when viewing duration is dictated by the environment. *Journal of Neuroscience*. *28*. 3017–3029 https://doi.org/10.1523/JNEUROSCI.4761-07.2008

Ossmy O., Moran R., Pfeffer T., Tsetsos K., Usher M., Donner T. H. (2013). The timescale of perceptual evidence integration can be adapted to the environment. *Current Biology*. *23*. 981–986. https://doi.org/10.1016/j.cub.2013.04.039

Piet A. T., El Hady A., Brody C. D. (2018). Rats adopt the optimal timescale for evidence integration in a dynamic environment. *Nature Communications. 9*. 4265. https://doi.org/10.1038/s41467-018-06561-y

Radillo A. E., Veliz-Cuba A., Josić K., Kilpatrick Z. P. (2017). Evidence Accumulation and Change Rate Inference in Dynamic Environments. *Neural Computation. 29*. 1561–1610. https://doi.org/10.1162/NECO_a_00957

Scott B. B., Constantinople C. M., Akrami A., Hanks T. D., Brody C. D., Tank D. W. (2017) . Fronto-parietal Cortical Circuits Encode Accumulated Evidence with a Diversity of Timescales. *Neuron. 95*. 385–398.e5. https://doi.org/10.1016/j.neuron.2017.06.013

Shadlen M. N, Kiani R, Hanks TD, Churchland A. K. (2008). An intentional framework. Better than conscious.; 71–101. Available: https://books.google.com/books?hl=en&lr=&id=JQLg7L-Qto4C&oi=fnd&pg=PA71&dq=shandlen+decision+making+2008&ots=HwhRIR764A&sig=BxKVCixK6nd5prLosh81RSt_sPM

Squire R. F, Noudoost B, Schafer R. J, Moore T. (2013). Prefrontal contributions to visual selective attention. *Annual Review Neuroscience. 36*. 451–466. https://doi.org/10.1146/annurev-neuro-062111-150439

Thura D, Cisek P. (2016). Modulation of Premotor and Primary Motor Cortical Activity during Volitional Adjustments of Speed-Accuracy Trade-Offs. *Journal of Neuroscience. 36*. 938–956. https://doi.org/10.1523/JNEUROSCI.2230-15.2016

Wong K-F, Wang X-J. (2006). A recurrent network mechanism of time integration in perceptual decisions. *Journal of Neuroscience. 26*. 1314–1328. https://doi.org/10.1523/JNEUROSCI.3733-05.2006

Bodily boundaries and beyond:

Exploring the malleability of bodily self-consciousness

An interview with Bigna Lenggenhager

by Jasmine T. Ho & Raphaël Millière

Cite as: Lenggenhager, B., Ho, J. T. & Millière, R. (2020). Bodily boundaries and beyond: Exploring the malleability of bodily self-consciousness. An interview with Bigna Lenggenhager by Jasmine T. Ho and Raphaël Millière. *ALIUS Bulletin*, *4*, 82-105, https://doi.org/10.34700/2vez-qq65

Bigna Lenggenhager
bigna.lenggenhager@psychologie.uzh.ch
University of Zürich,
Zürich, Switzerland

Raphaël Millière
rm3799@columbia.edu
Columbia University,
New York, USA

Jasmine T. Ho
jasmine.ho@uzh.ch
University of Zürich,
Zürich, Switzerland

Abstract

In this interview, Bigna Lenggenhager discusses her groundbreaking empirical work on bodily self-consciousness, bodily disorders and bodily illusions. The conversation explores issues related to the interpretation of the rubber hand illusion and the full-body illusion, the nature of the relationship between self-consciousness and bodily awareness, syndromes of disembodiment, as well as the use of virtual reality as a therapeutic tool for bodily disorders.

keywords: *self-consciousness, bodily awareness, rubber hand illusion, full-body illusion, virtual reality*

What sparked your interest in bodily self-consciousness? What types of questions were you aiming to answer by entering this field?

As a child, I read, or rather tried to read, many of my father's books on yoga, meditation, and Buddhism. Although I didn't understand much, I was fascinated by the different understanding of the self, life, and death in these traditions. Later during high school, my desire grew to study the human brain in its complexity, hoping to find answers to some of the big questions about how we perceive, think, and act, and why we are who we are. I went to university to study psychology, psychopathology, and neurophysiology.

During this time, I got interested in illusions and their value in revealing how the brain works. I was intrigued by visual and multisensory illusions, but above all by bodily illusions, such as the rubber hand illusion (Botvinick & Cohen, 1998) and Pinocchio illusion (Lackner, 1988). I found it fascinating how easily the perception of the most familiar object, i.e., one's own body, can be altered by just using a clever experimental design and without any drugs or long meditation techniques. But definitively the strongest driver of my interest in bodily self-consciousness were the patients I investigated during my first internship in the Neurology Department of the University Hospital of Zurich. Listening to their stories made me want to understand everything about how we perceive our body and how it links to our sense of self and consciousness. Realizing how many aspects of the perception of the self can be altered, I wanted to learn why and how we normally perceive our body (but not another person's body) as belonging to us, how we feel in control of our body, and how we perceive and act from a physically embodied perspective.

There are many disorders of bodily self-consciousness. Is there a specific condition that elicits a particular fascination for you?

From the external perspective as a researcher, all disorders of bodily self-consciousness are extremely fascinating to me. It is impressive how many different aspects of bodily self-consciousness can be affected. I think this broad range of symptoms makes bodily self-consciousness disorders so fascinating, ranging from the feeling of being duplicated or even multiplied, as in a specific form of heautoscopy (Brugger et al., 2006), or the feeling that one does not exist, as in Cotard syndrome.

What's more, different disorders fascinate me for different reasons. Take for example, out-of- body experiences: I find it extremely fascinating that these experiences have been described in so many different contexts and cultures, and that they might actually have shaped the way people think about the relation between the body and the mind, and between life and death. And the fact that such complex illusions can be induced by local brain stimulation (Blanke et al., 2002) remains fascinating to me.

Another example is Body Integrity Dysphoria, a disorder in which otherwise healthy individuals feel like a part of their body does not belong to them (Brugger et al., 2013). Here, the most fascinating aspect for me is probably the related societal and ethical questions, which heavily depend on our conception of how body, brain, and mind are linked.

In 2007, you published the first experimental study investigating the so-called full-body illusion (Lenggenhager et al., 2007). Since this landmark publication, you have done pioneering work on various versions of this illusion that have yielded a wealth of insights on the multisensory neural mechanisms underlying bodily self-consciousness in humans. Could you describe the basic setup of the full-body illusion, and summarize the main insights that you have derived from your work on this illusion?

In the mentioned study, we tried to bridge the knowledge and theories about the fundamental mechanisms of out-of-body experience (Blanke et al., 2005) with at that time recent insights from the plasticity of the bodily self in healthy participants using multisensory stimulation paradigms. We tried to extend the technique used in the famous rubber hand illusion paradigm (Botvinick & Cohen, 1998) using 3D video-based virtual reality systems to create more autoscopic or out-of-body-like illusions. We had great fun in the lab exploring and trying out many versions of such illusions.

In the final setup, participants were shown their own (or an object's or a mannequin's) body projected two meters in front of them. Then, by touching their back while displaying it synchronously (as compared to asynchronously) on the body projection, we enhanced self-identification with their projected body, which also led to a change in perceived self-location.

For me personally, the most interesting finding of this study was that we not only altered the perception of a single body part, as during the rubber hand illusion, but that even the perception of the whole body and self can be changed. This is interesting from a theoretical perspective but might also spark potential therapeutic applications (see e.g. (Pamment & Aspell, 2017)).

In a commentary from 2010, Adrian Alsmith writes the following about full-body illusions:

> "Unfortunately, the question of whether or not there are full-body illusions is empirically under-determined, as putative full-body illusions are difficult to isolate from illusions involving composite parts that do not constitute a 'full' or 'whole' body. That is to say, a plausible alternative is that only representations of the body parts directly stimulated become subject to the experimentally induced bias, whilst other parts remain relatively (perhaps even completely) unaffected." (Smith, 2010)

In other words, Alsmith suggests that the so-called full-body illusion might in fact be something like a "trunk illusion" in the classic setup. Do you think this is plausible? How do you think this question could be empirically settled? For example, if multiple body parts were stroked instead of just the trunk, would you expect to see a change in the effect size of implicit measurements and questionnaire reports?

I agree with Alsmith that with our experimental setup, we cannot differentiate between these two alternative explanations. We asked the participants in the self-location task to go back to "where they were standing before". This question/measure does not allow us to localize the trunk in a different place than let's say the arm. I think it would be rather easy to experimentally test Alsmith's hypothesis, both by stimulating more and different body parts (as in (Salomon et al., 2013)) or by using more sophisticated self-location measures in virtual reality that could be easily adapted to different body parts (Nakul et al., 2020). However, I do not think that his alternative explanation is very likely. Firstly, none of the participants has ever reported a sense of disruption or fragmentation of the body (which would, of course, be very interesting but does not seem very common in clinical conditions either). Secondly, a study of Ehrsson's group (Gentile et al., 2015) looked at generalizability of ownership from the stroked body part to the full body. Even if their setup was slightly different – with a full body illusion from the first-person perspective – the results clearly show that the sense of ownership spread to the non-stimulated body parts.

> "We not only altered the perception of a single body part, as during the rubber hand illusion, but even the perception of the whole body and self can be changed."

The rubber hand illusion (Botvinick & Cohen, 1998) represents the pioneering and most widely used experimental paradigm to temporarily alter body ownership by transferring the sense of ownership from one's own arm to an artificial body part. Synchronous stroking of the participant's concealed real hand and visible rubber hand, so that the felt touch of the brush on the real hand and the seen touch on the fake hand are closely matched, induces a sense of ownership of the artificial limb. Broadly accepted theoretical frameworks maintain that the rubber hand illusion reflects the role of multimodal integration in embodiment (Suzuki et al., 2013), indicative of the brain's attempt to rectify discrepant multisensory visuotactile information.

However, recent criticism contends that existing interpretations overlook the role of trait differences in the ability to generate experiences that meet expectancies (phenomenological control) (Lush, 2020). Demand characteristics, specifically participant expectancies and stable trait suggestibility, could account for the rubber hand illusion experiences by "generating expectancies which are met by the voluntary top-down control of phenomenology" (Lush, 2020, p. 1), similar to imaginative suggestibility within the context of hypnosis. This line of argumentation rests on empirical results evincing that measures of the rubber hand illusion are substantially related to hypnotic and sensory suggestibility (Fiorio et al., 2020; Marotta et al., 2016; Walsh et al., 2015), and that expectancies predict illusion scores (Lush et al., 2019).

Do you think that measures of the rubber hand illusion may be confounded by the active generation of phenomenological control? If a lack of control of demand characteristics truly exists, how could rubber hand illusion measures be improved? Would the aforementioned criticisms extend to full body illusions as well, particularly those completed in virtual reality, which would not entail contradictory proprioceptive information between the real and artificial body part?

Yes, I think it is certainly worth it and important to follow up this line of research and alternative explanations in more detail and with solid experimental methods. Other than in classical visual illusions, like the Müller-Lyer illusion (Müller-Lyer, 1889), which are perceived by everyone and extremely stable between participants, in bodily illusions we typically observe strong individual differences in how they are perceived and how strongly people react to them. On the other hand, when you observe how people react to walking over a virtual plank in the air, like in Vive's plank

simulation, you can tell that even if they have insight into the experimental setup, the body reacts as if they didn't. I agree that individual differences could partially be caused by differences in suggestibility, and I find it surprising myself how little research directly addresses why people react so differently and what potential underlying mechanisms might be at work. I think it is generally important to put more effort in understanding differences between various individuals and also between various settings, but do not think that demand characteristics can explain everything. Sebastian Dieguez recently suggested, in a critical comment to one of our papers, to call bodily illusions rather "aliefs" (Dieguez, 2018). He suggests that such aliefs depend on motivation and expectancy of the participants, which is very much in line with Lush's findings. But I think, even if future empirical work would indeed show that demand characteristics are the or an explanation of bodily illusions, it does not make the findings of the bodily self literature less interesting (Roel Lesur et al., 2018). Over the last 20 years, a vast amount of studies have shown that with bodily illusions, many different physiological, emotional, and cognitive aspects can be altered. For example, hippocampal activity underlying autobiographical memory (Bergouignan et al., 2014), implicit peripersonal space measures (Noel et al., 2015), and immunological response (Barnsley et al., 2011) have shown to be modulated by bodily illusions. If these processes can all be modified by phenomenological control in the predicted way, this would seem quite impressive to me. Together with you, Jasmine, we are currently investigating the link between embodiment, illusory embodiment, and placebo response, which I think is a promising research path to continue. A further aspect that I think might be difficult to explain with phenomenological control are the increasing findings in animals on "bodily illusion", suggesting that there are physiological changes specifically to the illusion condition (e.g., Shokur et al., 2013; Wada et al., 2016). This is, by the way, another line of research that I think would be exciting to follow up. For example: how does firing rate of place or grid cells change after a full body illusion in rats? This would be a much more implicit measure of self location than the one we used in our first study you mentioned above.

Regarding the question on how we can improve current research, I think it would be important to develop more implicit measures, but also (try) to

replicate the many interesting implicit measures in humans and animals that have already been described, ideally in much bigger sample sizes. This is important, as some of these beautiful data have not yet or only rarely been replicated. Bigger sample size would facilitate the assessment of individual differences and to link different implicit measures with phenomenological changes.

To follow up on the points above, the rubber hand illusion is classically considered a measure of body ownership. Yet, accumulating evidence suggests that body ownership and self-location constitute distinct aspects of bodily self-consciousness (Serino et al., 2013). While explicit measures of body ownership are assessed by means of a self-report questionnaire, proprioceptive drift towards the rubber hand is generally considered an implicit measure of body ownership. However, if proprioception measures the shift of the own arm towards the rubber hand, would implicit proprioceptive drift not more accurately constitute a representation of self-location rather than ownership? Considering that these implicit and explicit measures of the rubber hand illusion further seem to be substantially yet differentially related to hypnotic and sensory suggestibility, would a reform or at least more specific operationalization of the rubber hand illusion measures be warranted?

Yes, I would agree. A measure that might be more directly linked to the sense of body ownership than the proprioceptive drift might be, at least conceptually, the electrodermal activity to threat (Armel & Ramachandran, 2003). However, there are also some problems with this method and we often find diverging results between threat measures and self-report questionnaires (e.g., Roel Lesur, Weijs, et al., 2020). As mentioned above, I think it will be important to relate the different measures of the rubber hand illusion to different phenomenological measures and also to potential confounding differences (e.g., the aforementioned suggestibility). For that, I think it will also be important to further develop the explicit measures we use, and I am very optimistic that some of the measures you, Raphaël, apply, like microphenomenological interviews combined with EEG based biomarkers (Timmermann et al., 2019) might advance this research direction.

The notion of "first-person perspective" (1PP) is widely used in the scientific literature on the full-body illusion and virtual reality (e.g., Blanke & Metzinger, 2009). However, its definition is not always consistent across publications. Here are a few examples:

> (a) “the feeling from where 'I' experience the world around me” (Pfeiffer et al., 2014; see also Ionta et al., 2011; Blanke, 2012).
>
> (b) “the point from which visual information from the environment is gathered” (Maselli, 2015).
>
> (c) “the experience of taking a first-person, body-centered, perspective on [one's] environment” (Serino et al., 2013).
>
> (d) “a purely geometrical feature of an egocentric model of reality [which] includes a spatial frame of reference, plus a global body representation, with a perspective originating within this body representation” (Blanke & Metzinger, 2009)
>
> (e) “the experience of being a subject and of being directed at the word” (Pfeiffer et al., 2016).

Many of these definitions (e.g., a, c and e) explicitly describe 1PP as an experience, but some (e.g., b and d) do not. Furthermore, some definitions include a reference to bodily representation, while others merely refer to the information conveyed by visual experience.

Beyond these lexical discrepancies, it seems that there is a broad difference between two ways of operationalizing the notion of 1PP. Some studies seem to construe 1PP as the subject's awareness of the egocentric location of the point of origin of her visuospatial spatial perspective – i.e., the location from where the subject sees their environment. Other studies seem to construe 1PP as the subject's awareness of the orientation of their head or body with respect to a geocentric frame of reference – e.g., whether the subject is looking up or down with respect to gravitational cues.

Presumably, these two constructs are not equivalent. Blind individuals (or sighted individuals wearing a blindfold) lack a visual perspective, but can nonetheless be aware of their orientation with respect to gravitational cues. Conversely, sighted individuals can become temporarily disoriented (for example because of vertigo, zero G, or galvanic vestibular stimulation), while remaining aware of the location from where they see their environment.

Do you think the notion of 1PP should be disambiguated, such that more specific constructs can be investigated empirically, or do you think the apparent discrepancies between construals of the notion can be reconciled in a single construct?

It is true that there are many different definitions and conceptions of the first-person perspective. Defining the first-person perspective as a visuo-spatial perspective, in my opinion, neglects our multisensory perception of the world, the body, and ourselves. This, I think, reflects the classical very strong dominance of vision and visual research in science. Thus, I would argue that at the very least, it should be called and considered as an audio-visuospatial perspective, since we clearly have an auditory egocentric perspective as well. It can be experimentally manipulated as well, similar to the visual one (e.g., Lesur et al., 2020) and can also be mismatched in rare clinical conditions (Blanke et al., 2004). Furthermore, as you mention, the gravitational perspective is also important and often neglected. But rather than disambiguate the notion of the first-person perspective, I think it is important to consider that in normal waking consciousness, our first-person perspective is based on the integration of these different perspectives. I think it is interesting and important to experimentally take them apart and investigate how they affect various other aspects of self-consciousness, such as for example self-location. To give an example, let's say the auditory perspective is different from the visual, where would you feel localized? At either one or the other perspective, alternating, bilocated, or somewhere in between?

Similarly, one might wonder whether the notion of "self-location" – also widely used in the scientific literature on the full-body illusion and virtual reality – could benefit from disambiguation. Self-location is often defined as "the experience of where I am in space" (e.g., Ionta et al., 2011; Blanke, 2012; Blanke et al., 2015; Pfeiffer et al., 2014).

In the full-body illusion with front stroking, three notions of self-location can potentially be distinguished:

1. The subject's awareness of the location of the point of origin of her visuospatial perspective in an egocentric frame of reference.

2. The subject's awareness of the location of her own body – that she sees in front of herself – in an egocentric frame of reference.

3. The subject's awareness of her location in an allocentric frame of reference (i.e., with respect to a map-like representation of her environment).

The first notion seems equivalent to one of the two notions of 1PP discussed in the previous section. Consequently, there might be some contexts in which "self-location" and "1PP" refer to a single construct. However, forward proprioceptive drift, which is used in many studies on the FBI as an implicit measure of self-location, appears to be more related to the third notion. Indeed, this measure is obtained by moving participants around the room with a blindfold, and asking them to walk back to their original location. Presumably, this requires subjects to represent their current location and their original location within an allocentric frame of reference, compute a path from one to the other, and translate the relevant coordinates in an egocentric frame of reference to carry out the adequate movements.

By contrast, it is perhaps more difficult to say which notion of self-location is probed by the mental ball-dropping task, which is used as an alternative implicit measurement of self-location in horizontal versions of the full body illusion. This task involves asking subjects to imagine that they are dropping a ball from their location, and to estimate how long the ball would take to hit the floor, as a way to assess how elevated the subjects take their location to be compared to ground level. In so far as one normally drops a ball with one's hand – as opposed to one's mouth, for example – subjects might use the second notion of self-location as their reference point for the mental ball-dropping task. Furthermore, the computations required to perform the task might not require translation of the relevant coordinates (that of the ball and of the ground level) into an allocentric frame of reference. Indeed, subjects should be able to estimate the egocentric distances of the fictional ball and of the ground level with respect to the point of origin of their visuospatial perspective, and carry out the mental task entirely in an egocentric frame of reference.

Do you think these brief comments about the targets of the implicit measurements used in the Full Body Illusion literature are plausible and accurate? Specifically, do you think the proprioceptive drift task and the mental ball-dropping task measure the same notion of "self-location"? If not, do you have suggestions about how to disambiguate between different notions of self-location in empirical studies using the Full Body Illusion?

Yes, I would agree with your interpretation of the different definitions and measures of self-location. As mentioned above, in normal waking consciousness, the multisensory first-person perspective and self-location are typically co-located, and should also be comparable in the different implicit

and explicit measurements of self-location, even if they require, as you mention, different mental processing (note that in the walking task, memory processes might also play a more important role). However, it might well be that, if you modulate one or the other perspective, be it the visual, the auditory, or the gravitational, this would affect the different measures differently. One step in this direction would be to use different assessments of self-location measures in the full body and other illusions, and relate them also to the subjective perception during different illusions, which is something we have recently started in collaboration with Christophe Lopez (Nakul et al., 2020). As mentioned above, I am in general not yet very happy with the self-location measures we are using. More implicit measures such as place cell activity would be helpful.

Although the topic surrounding the self-awareness of animals remains beset with a number of difficulties, a few studies have demonstrated that non-human primates and mice may experience body ownership as well. These illusions similarly utilize multisensory stimulations akin to the rubber hand illusion in humans (i.e., (a)synchronous stimulation of a physical and artificial arm or tail). Do you think such bodily illusions can effectively translate to animal studies and validly allude to the presence of bodily self-consciousness in animals? What type of paradigm would you employ if you were to conduct such a study?

If you see how much difficulty healthy adult participants have in describing what exactly they feel during bodily illusions, I think it is very hard to imagine what animals would feel during such illusions and even harder to imagine how you would measure such or quantify these feelings. Nevertheless, as mentioned above, I think this literature is interesting and important for various reasons (e.g., the phenomenological control). If maybe not telling us much about bodily self-consciousness, animal studies might certainly help to shed light on basic underlying mechanisms, especially on multisensory weighing and integration. Next to the full body illusion in rats that I mentioned above, I would really love to do longer term experiments in, e.g., mice, if it is possible to do it in an ethically acceptable way, to better understand the plasticity of such multisensory integration mechanisms, which you cannot do in humans. What if a rat for example would, using binaural headphones, from birth on always hear from another rat's perspective (or from a stationary perspective in the room similar to (Mizumoto & Ishikawa, 2005)), while seeing and smelling from their own

perspective? How would that influence the rat's perceived self-location, as for example measured with a reaction to threat on various positions? Investigating such long-term multisensory mismatching stimulation would certainly tell us a lot about the plasticity of multisensory mechanisms underlying the bodily self.

Several influential theories presuppose that basic, largely implicit and pre-reflective bodily processes underlie self-consciousness, where the integration of sensory and motor bodily signals with the self as an agent of intentional object remains anchored in an embodied self (Blanke & Metzinger, 2009; S. Gallagher, 2005, 2013; Newen, 2018). On the contrary, opponents of this claim do not consider bodily sensations or components of bodily awareness (i.e., sense of ownership, self-location, and agency) a necessary prerequisite for self-conscious experience, or even an essential requirement for consciousness in general (Millière, 2020; Millière & Metzinger, 2020).

Cases from clinical samples encompass unique aberrations from ordinary bodily self-consciousness; for example, patients with asomatognosia suffer from an unawareness of ownership of a part of their body (loss of ownership), patients with tetraplegia experience paralysis that can result in the total loss of movement and sensation in all four limbs and torso (loss of agency), whereas out-of-body experiences from neurological origins include the feeling of disembodiment from one's own body and viewing it from an elevated visuospatial perspective (alterations of self-location). Individuals with body integrity dysphoria may not feel like a limb belongs to them and desire amputation or being paraplegic. Accumulating empirical evidence suggests that transient absence of bodily experiences or weakly embodied states are also encountered in non-clinical populations. Examples of such states include feelings of disembodiment or a detachment between the body and mental processes under the influence of psychedelic substances (Preller & Vollenweider, 2018; Timmermann et al., 2019; Vollenweider & Kometer, 2010) or a loss of spatiotemporal awareness in dreams (Occhionero & Cicogna, 2011).

Do you think that such experiences contradict the claims that purport a putative role of the body for self-consciousness? One might argue that even in the presence of loss of agency (e.g., tetraplegia), alterations of self-location (e.g., out-of-body experiences), or loss of self-location (reported in some drug-induced states, in sensory deprivation and in deaf blindness, (Millière, 2019)) one could nevertheless preserve a connection to a physical body and therefore to bodily self-consciousness. In the presence of "bodiless" dreams,

however, one may retain a sense of agency, self-location, and first-person perspective, despite lacking the experience of a body. Would it still make sense to strap such states to bodily self-consciousness, or would it be more appropriate to talk of perspectival consciousness?

I think that the mentioned examples do not contradict the claim of an important role of the body in self-consciousness for several reasons. First of all, I would argue that even if we find some rare instances (for example pathological or drug-induced) in which self-consciousness might be possible without momentary bodily experiences, it does not at all mean that bodily self-consciousness is not generally important in self-consciousness. There is the famous saying that the exception confirms the rule, and we know that the brain is highly plastic and adaptive, thus it is difficult to generalize from very specific cases to "normal" consciousness. Furthermore, all examples given are specific states in individuals who have or at least had experienced normal bodily self-awareness previously. For example, a patient with somatoparaphrenia might lose the sense of ownership for one hand but still has a normal sense of ownership for the other hand. A tetraplegic might still have a normal sense of agency over some body parts like the eye-lids. And a dreaming person senses her body in waking state. These instances might be an important prerequisite to be able to transiently perceive a "bodiless" state.

Thus, I would still see the lack of the sense of a body as a variation of an altered body perception. After all, even if you wanted to call it perspectival consciousness, a visual or auditory perspective is only experienced the way we experience it, due to the way our body is physically shaped.

> “ I would still see the lack of the body as a variation of an altered body perception ”

One can certainly argue that persons with surgical or congenital amputations retain a sense of bodily self-consciousness, despite the lack or loss of their limbs. In the most severe cases, individuals may physically comprise no more than their trunk and head. Yet, the fundamental sense of selfhood associated with bodily self-consciousness is ostensibly experienced as a global whole-body representation, rather than individual representation of separate body parts (Lenggenhager et al., 2007). In theory – and not considering the constraints of vital organs – how much of the body could we progressively

remove before the sense of (bodily) self-consciousness is lost? Along these lines, could prosthetics serve as a viable substitute for the physical body?

This is an intriguing question. Similar to the famous ship of Theseus example, which questions whether a ship would still be the same if all parts were successively replaced over time. This example would be even more interesting for a human being than for a ship, what if you replaced one organ/body part after the other in a human being, would she still be the same? While this remains thus far a thought experiment, the question on how self-consciousness is altered by, e.g., organ transplantation or prosthetics, is a very interesting and important one, which should be carefully investigated. I think the "problem" in your question might exactly be the "not considering the constraints of vital organs", as both empirical and theoretical evidence suggest that the trunk and head, where also the vital organs are located, is the core of this fundamental selfhood, which is also where people typically localize their "self" if forced to localized it in a single point (Alsmith & Longo, 2014). This also fits the idea that interoceptive cues, mainly from the trunk, might crucially underlie our sense of a bodily self (e.g., Park & Blanke, 2019). But in a pure thought experiment, my guess would be that as long as you have some physical body (even a prosthetic one if it were integrated in your sensorimotor loop) you would feel as physically embodied. And again, I would argue that there is an important difference between whether you previously had a body and lose/replace it as in your thought experiment, or whether there would be no body from the beginning.

Several clinical trials using psychedelic substances have evinced impressive and often persistent improvements in depression. During such sessions, perceived disembodiment or detachment from the physical body constitute common phenomenological experiences under the influence of psychedelic substances (Belser et al., 2017; Watts et al., 2017). In a separate line of research, an experimentally induced out-of-body illusion successfully alleviated symptoms in a number of chronic pain conditions, including fibromyalgia, endometriosis, chronic lower back pain, and spinal cord injury (Pamment & Aspell, 2017). Do you think that altered states of bodily self-consciousness, specifically those related to disembodiment, are central to the alleviation of symptoms in such conditions? If so, why and how would a perceived "detachment" from our bodies operate beneficially for our mental and physical health?

While the mentioned examples are very promising there is still limited empirical evidence. Replication studies are needed. However, I think that the experience of disembodiment, as well as possibly related sensations, such as the sense of lightness, might be beneficial. Again, I think it would be important to better quantify different phenomenological aspects of such altered sense of embodiment and to investigate in the different states (e.g., in virtual reality or during psychedelic drugs) to better understand which aspects of such experiences might be helpful. Generally, I am very excited by the idea that the unusual body experience itself rather than, for example, the pure neurochemical alterations due to psychedelic substances might cause some of these changes. If you think of the hyperembodiment model of Thomas Fuchs, for example (Fuchs & Schlimme, 2009), in which depression is suggested to be linked to a too rigid and strong sense of embodiment, it could well be that the sense of transient disembodiment or lightness might be helpful. Furthermore, while a constant sense of detachment from the body might not be beneficial, I think that there is a lot of therapeutic potential in using such tools to let people perceive how plastic the sense of the body actually is. This could also be applied in a more educative way by showing people that the sense of their own body might not be as rigid as they believe.

“While a constant sense of detachment from the body might not be beneficial, I think that there is a lot of therapeutic potential in using such tools to let people perceive how plastic the sense of the body actually is”

Virtual reality has undergone a transition from an expensive and arduous device to a functional technology that is increasingly employed both in empirical and home-use settings. While the immersive nature of virtual reality has already demonstrated considerable success as an application for behavioral health (e.g., exposure therapy for phobias or posttraumatic stress disorder) (Riva et al., 2018), a particularly auspicious feature of virtual reality involves the implementation of virtual avatars as bottom-up sensory modulators of existing body representations (embodied virtual reality). Virtual reality permits a degree of sensory control that would not be possible in physical reality, and thus facilitates unique opportunities to update aberrant representations of bodily self-consciousness.

With a potentially increasing recognition of embodiment in the convergence of phenomenology and pathology, do you consider embodied virtual reality a promising therapeutic tool for disorders of bodily self-consciousness? Is it or will it be possible to effectively "replace" our physical bodies with virtual selves? Given adequate sensory information, can we embody almost anything – e.g., an animal, a superhero, or an inanimate object such as a house? Along these lines, could we further embody multiple selves, and if so, how might this affect our sense of (bodily) self long term? Although the potential applications for clinical disorders seem promising, what are some of the ethical concerns and risks that need to be considered moving forward with virtual reality and other "body surrogates", such as robotics and prosthetics?

Yes, I definitely consider embodied virtual reality a promising therapeutic tool and potentially helpful for educational settings or training situations. We have recently shown that people can even feel like they are strongly embodying a grapefruit (Lesur, Aicher, et al., 2020). That's why I am optimistic that healthy participants can basically embody anything, at least as long as head-related visuo-motor coherency is given (Lesur et al., 2018b), even if it might be an alief rather than an illusion. While I have never tried to virtually embody several bodies and have a hard time imagining it, previous literature suggests it is possible (e.g., Heydrich et al., 2013). I agree with you that it is important to do more basic research before developing too many therapeutic applications, especially when it comes to children, who are presumably still developing their sense of a bodily self. I am surprised how little is known about how embodied virtual reality exposure alters (bodily) self-consciousness in children. It is one of my core objectives of my current Swiss National Science Foundation-funded project to investigate these mechanisms. It is important to carefully evaluate ethical considerations (e.g., Madary & Metzinger, 2016), especially as potential therapeutic embodied virtual reality, as well as embodied virtual reality for leisure (e.g., in games), will use much longer exposure times than what we typically use in the laboratory.

In the virtual reality literature, it is often claimed that good virtual reality technology should induce a "sense of presence", defined as the sense of being present within the virtual environment that one perceived through a head-mounted display (e.g., Heeter, 1992; Held & Durlach, 1992; Slater, 2009).

There are a number of technical specifications of virtual reality systems that appear to be strong mediators, if not requirements, for the induction of a sense of presence. These include, among others: stereoscopy (providing binocular depth cues), head tracking (providing action-contingent visual feedback), hand/body tracking (giving the user a virtual body), high screen refresh rate (enabling smooth visual motion), high screen resolution (providing clear visual input), and wide field of view (providing peripheral visual input). A meta-analysis of the relationship between the sense of presence (as measured by various questionnaires) and different technical specifications virtual reality systems in 83 studies found that head tracking was one of the most important features to induce a sense of presence (Cummings & Bailenson, 2016).

When these technical specifications are not met, the experience of virtual reality users can become quite uncomfortable. In particular, the lack of head tracking (or head tracking with high latency) causes a mismatch between visual input and proprioceptive/vestibular cues about self-motion, since the viewpoint rendered by the computer is not sensitive to the user's head movements. Such mismatch is often associated with an unpleasant combination of symptoms known as 'cybersickness' – a special case of motion sickness (Gallagher & Ferrè, 2018). There is some evidence that cybersickness is negatively correlated with scores on various presence questionnaires (Weech et al., 2019). This raises the following question: is the sense of presence in virtual reality really a positive experience that one has – in addition to whatever else one might experience while using virtual reality –, or is it simply the absence of the discomfort and abnormal sensations, such as cybersickness, associated with less sophisticated virtual reality systems? Do you think one of these two hypotheses is more plausible than the other, on the basis of evidence provided by your own research or virtual reality research at large?

I think it depends on the content of the virtual reality. If you imagine a virtual reality that exactly simulates reality, I would clearly expect that a strong sense of presence is just the absence of any discomfort and not per se a pleasant sensation. However, if you take virtual reality as a tool to let participants feel the sense of presence in a pleasant and surprising world that they otherwise could not be present, apart from maybe in dreams or during a psychedelic experience or similar, the mere sense of presence might be associated with a positive experience. But of course, depending on the content of the environment this could also rapidly turn into a nightmare.

If ethics, technology, and finances were no concern, what would be your dream research study?

Hmm, as you know I am a great fan of the virtual body swapping, as for example in the Machine To Be Another (Oliveira et al., 2016) and I love trying out all experiments as a participant before I start an experiment. My dream study would be to swap bodies in reality rather than in virtual reality. Maybe related to what was mentioned before, to create an experiment in which you could progressively swap out body parts and organs of the other person and test how it changes affective and cognitive processes, especially the sense of a bodily self and self-consciousness. If this answer is a bit too science-fictional, I would dream to do a study in humans as I suggested in animals above, letting them grow up with permanently altered multisensory perspectives or multisensory contingencies. If that sounds too scary, I would at least love to do more long-term exposure experiments. But even those need a lot of money and facilities. And there might of course still be ethical concerns.

With the rapid advancement of technology, what types of questions do you anticipate the field of bodily self-consciousness research will be attempting to answer in 20 years?

The question on how digitalization and digital interactions change bodily self-consciousness will be increasingly relevant. The fact that we interact more and more in quasi disembodied states through various digital media might shape the way we perceive ourselves and others. The fact that we might increasingly have the possibility to digitally change and represent certain aspects of our body and self in much more flexible ways (e.g., using facial filters, self-chosen avatars or holograms) might change our notion of self-consciousness and reveal new ethical questions.

“ The fact that we interact more and more in quasi-disembodied states through various digital media might shape the way we perceive ourselves and others. ”

References

Alsmith, A., & Longo, M. R. (2014). Where exactly am I? Self-location judgements distribute between head and torso. *Consciousness and Cognition, 24*, 70– 74. https://doi.org/10.1016/j.concog.2013.12.005

Armel, K. C., & Ramachandran, V. S. (2003). Projecting sensations to external objects: Evidence from skin conductance response. *Proceedings of the Royal Society of London B: Biological Sciences, 270*(1523), 1499–1506. https://doi.org/10.1098/rspb.2003.2364

Barnsley, N., McAuley, J. H., Mohan, R., Dey, A., Thomas, P., & Moseley, G. L. (2011). The rubber hand illusion increases histamine reactivity in the real arm. *Current Biology, 21*(23), R945–R946. https://doi.org/10.1016/j.cub.2011.10.039

Belser, A. B., Agin-Liebes, G., Swift, T. C., Terrana, S., Devenot, N., Friedman, H. L., Guss, J., Bossis, A., & Ross, S. (2017). Patient Experiences of Psilocybin-Assisted Psychotherapy: An Interpretative Phenomenological Analysis. *Journal of Humanistic Psychology, 57*(4), 354–388. https://doi.org/10.1177/0022167817706884

Bergouignan, L., Nyberg, L., & Ehrsson, H. H. (2014). Out-of-body–induced hippocampal amnesia. *Proceedings of the National Academy of Sciences, 111*(12), 4421–4426. https://doi.org/10.1073/pnas.1318801111

Blanke, O. (2012). Multisensory brain mechanisms of bodily self-consciousness. Nature Reviews Neuroscience, 13(8), 556–571. https://doi.org/10.1038/nrn3292

Blanke, O., Landis, T., Spinelli, L., & Seeck, M. (2004). Out-of-body experience and autoscopy of neurological origin. *Brain, 127*(2), 243–258. https://doi.org/10.1093/brain/awh040

Blanke, O., & Metzinger, T. (2009). Full-body illusions and minimal phenomenal selfhood. *Trends in Cognitive Sciences, 13*(1), 7–13. https://doi.org/10.1016/j.tics.2008.10.003

Blanke, O., Mohr, C., Michel, C. M., Pascual-Leone, A., Brugger, P., Seeck, M., Landis, T., & Thut, G. (2005). Linking Out-of-Body Experience and Self Processing to Mental Own-Body Imagery at the Temporoparietal Junction. *Journal of Neuroscience, 25*(3), 550–557. https://doi.org/10.1523/JNEUROSCI.2612-04.2005

Blanke, O., Ortigue, S., Landis, T., & Seeck, M. (2002). Stimulating illusory own-body perceptions. *Nature, 419*(6904), 269–270. https://doi.org/10.1038/419269a

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., Blanke, O., Regard, M., Bradford, D. T., & Landis, T. (2006). Polyopic Heautoscopy: Case Report and Review of the Literature. *Cortex, 42*(5), 666–674. https://doi.org/10.1016/S0010-9452(08)70403-9

Brugger, P., Lenggenhager, B., & Giummarra, M. J. (2013). Xenomelia: A Social Neuroscience View of Altered Bodily Self-Consciousness. *Frontiers in Psychology, 4*. https://doi.org/10.3389/fpsyg.2013.00204

Cummings, J. J., & Bailenson, J. N. (2016). How Immersive Is Enough? A Meta-Analysis of the Effect of Immersive Technology on User Presence. *Media Psychology, 19*(2), 272–309. https://doi.org/10.1080/15213269.2015.1015740

Dieguez, S. (2018). The Illusion Illusion. Constructivist Foundations, 14(1), 108–110.

Fiorio, M., Modenese, M., & Cesari, P. (2020). The rubber hand illusion in hypnosis provides new insights into the sense of body ownership. *Scientific Reports, 10*(1), 5706. https://doi.org/10.1038/s41598-020-62745-x

Fuchs, T., & Schlimme, J. E. (2009). Embodiment and psychopathology: A phenomenological perspective. *Current Opinion in Psychiatry, 22*(6), 570–575. https://doi.org/10.1097/YCO.0b013e3283318e5c

Gallagher, M., & Ferrè, E. R. (2018). Cybersickness: A Multisensory Integration Perspective. *Multisensory Research, 31*(7), 645–674. https://doi.org/10.1163/22134808-20181293

Gallagher, S. (2005). How the Body Shapes the Mind. *Oxford University Press*. http://www.oxfordscholarship.com/view/10.1093/0199271941.001.0001/acprof-9780199271948

Gallagher, S. (2013). A Pattern Theory of Self. *Frontiers in Human Neuroscience, 7*. https://doi.org/10.3389/fnhum.2013.00443

Gentile, G., Björnsdotter, M., Petkova, V. I., Abdulkarim, Z., & Ehrsson, H. H. (2015). Patterns of neural activity in the human ventral premotor cortex reflect a whole-body multisensory percept. *NeuroImage, 109*, 328–340. https://doi.org/10.1016/j.neuroimage.2015.01.008

Heeter, C. (1992). Being There: The Subjective Experience of Presence. Presence: Teleoperators and Virtual Environments, 1(2), 262–271. https://doi.org/10.1162/pres.1992.1.2.262

Held, R. M., & Durlach, N. I. (1992). Telepresence. Presence: Teleoperators and Virtual Environments, 1(1), 109–112. https://doi.org/10.1162/pres.1992.1.1.109

Heydrich, L., Dodds, T., Aspell, J., Herbelin, B., Buelthoff, H., Mohler, B., & Blanke, O. (2013). Visual capture and the experience of having two bodies – Evidence from two different virtual reality techniques. *Frontiers in Psychology, 4*. https://doi.org/10.3389/fpsyg.2013.00946

Ionta, S., Heydrich, L., Lenggenhager, B., Mouthon, M., Fornari, E., Chapuis, D., Gassert, R., & Blanke, O. (2011). Multisensory Mechanisms in Temporo-Parietal Cortex Support Self-Location and First-Person Perspective. *Neuron, 70*(2), 363–374. https://doi.org/10.1016/j.neuron.2011.03.009

Lackner, J. R. (1988). Some proprioceptive influences on the perceptual representation of body shape and orientation. *Brain, 111*(2), 281–297. https://doi.org/10.1093/brain/111.2.281

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

Lesur, M. R., Bolt, E., & Lenggenhager, B. (2020). The monologue of the double: Allocentric reduplication of the own voice alters bodily self perception. BioRxiv, 2020.08.11.246397. https://doi.org/10.1101/2020.08.11.246397

Lush, P. (2020). Demand Characteristics Confound the Rubber Hand Illusion. *Collabra: Psychology, 6*(1), 22. https://doi.org/10.1525/collabra.325

Lush, P., Botan, V., Scott, R. B., Seth, A., Ward, J., & Dienes, Z. (2019). Phenomenological control: Response to imaginative suggestion predicts measures of mirror touch synaesthesia, vicarious pain and the rubber hand illusion. https://doi.org/10.31234/osf.io/82jav

Madary, M., & Metzinger, T. K. (2016). Real virtuality: a code of ethical conduct. Recommendations for good scientific practice and the consumers of VR-technology. Frontiers in Robotics and AI, 3, 3. https://doi.org/10.3389/frobt.2016.00003

Marotta, A., Tinazzi, M., Cavedini, C., Zampini, M., & Fiorio, M. (2016). Individual Differences in the Rubber Hand Illusion Are Related to Sensory Suggestibility. *PLOS ONE, 11*(12), e0168489. https://doi.org/10.1371/journal.pone.0168489

Maselli, A. (2015). Allocentric and egocentric manipulations of the sense of self-location in full-body illusions and their relation with the sense of body ownership. *Cognitive Processing, 16*(1), 309–312. https://doi.org/10.1007/s10339-015-0667-z

Millière, R. (2019). Are there degrees of self-consciousness? *Journal of Consciousness Studies, 26*(3–4), 252–276.

Millière, R. (2020). The varieties of selflessness. *Philosophy and the Mind Sciences, 1*(I), 8. https://doi.org/10.33735/phimisci.2020.I.48

Millière, R., & Metzinger, T. (2020). Radical disruptions of self-consciousness. *Philosophy and the Mind Sciences, 1*(I), 1–1. https://doi.org/10.33735/phimisci.2020.I.50

Mizumoto, M., & Ishikawa, M. (2005). Immunity to Error through Misidentification and the Bodily Illusion Experiment. *Journal of Consciousness Studies, 12*(7), 3–19.

Müller-Lyer, F. (1889). Optische Urteilstäuschungen, du Bois Arch. *Archiv Für Physiologie Suppl., 263*, 263–270.

Nakul, E., Orlando-Dessaints, N., Lenggenhager, B., & Lopez, C. (2020). Measuring perceived self-location in virtual reality. *Scientific Reports, 10*(1), 6802. https://doi.org/10.1038/s41598-020-63643-y

Newen, A. (2018). The Embodied Self, the Pattern Theory of Self, and the Predictive Mind. *Frontiers in Psychology, 9*. https://doi.org/10.3389/fpsyg.2018.02270

Noel, J.-P., Pfeiffer, C., Blanke, O., & Serino, A. (2015). Peripersonal space as the space of the bodily self. *Cognition, 144*, 49–57. https://doi.org/10.1016/j.cognition.2015.07.012

Occhionero, M., & Cicogna, P. C. (2011). Autoscopic phenomena and one's own body representation in dreams. *Consciousness and Cognition, 20*(4), 1009–1015. https://doi.org/10.1016/j.concog.2011.01.004

Oliveira, E. C. D., Bertrand, P., Lesur, M. E. R., Palomo, P., Demarzo, M., Cebolla, A., Baños, R., & Tori, R. (2016). Virtual Body Swap: A New Feasible Tool to Be Explored in Health and Education. 2016 XVIII Symposium on Virtual and Augmented Reality (SVR), 81–89. https://doi.org/10.1109/SVR.2016.23

Pamment, J., & Aspell, J. E. (2017). Putting pain out of mind with an 'out of body' illusion. *European Journal of Pain, 21*(2), 334–342. https://doi.org/10.1002/ejp.927

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

Pfeiffer, C., Grivaz, P., Herbelin, B., Serino, A., & Blanke, O. (2016). Visual gravity contributes to subjective first-person perspective. *Neuroscience of Consciousness, 2016*(1), niw006. https://doi.org/10.1093/nc/niw006

Pfeiffer, C., Schmutz, V., & Blanke, O. (2014). Visuospatial viewpoint manipulation during full-body illusion modulates subjective first-person perspective. *Experimental Brain Research, 232*(12), 4021–4033. https://doi.org/10.1007/s00221-014-4080-0

Preller, K. H., & Vollenweider, F. X. (2018). Phenomenology, Structure, and Dynamic of Psychedelic States. In A. L. Halberstadt, F. X. Vollenweider, & D. E. Nichols (Eds.), Behavioral Neurobiology of Psychedelic Drugs (pp. 221–256). Springer Berlin Heidelberg. https://doi.org/10.1007/7854_2016_459

Riva, G., Wiederhold, B. K., & Mantovani, F. (2018). Neuroscience of Virtual Reality: From Virtual Exposure to Embodied Medicine. *Cyberpsychology, Behavior, and Social Networking, 22*(1), 82–96. https://doi.org/10.1089/cyber.2017.29099.gri

Roel Lesur, M., Aicher, H., Delplanque, S., & Lenggenhager, B. (2020). Being Short, Sweet, and Sour: Congruent Visuo-Olfactory Stimulation Enhances Illusory Embodiment. *Perception, 49*(6), 693–696. https://doi.org/10.1177/0301006620928669

Roel Lesur, M., Gaebler, M., Bertrand, P., & Lenggenhager, B. (2018a). Authors' Response: On the Components and Future Experimental Setups of Bodily Illusions/Aliefs. *Constructivist Foundations, 14*(1), 111–113.

Roel Lesur, M., Gaebler, M., Bertrand, P., & Lenggenhager, B. (2018b). The Plasticity of the Bodily Self: Head Movements in Bodily Illusions and Their Relation to Gallagher's Body Image and Body Schema. Constructivist Foundations, 14(1), 94–105.

Roel Lesur, M., Weijs, M. L., Simon, C., Kannape, O. A., & Lenggenhager, B. (2020). Psychometrics of Disembodiment and Its Differential Modulation by Visuomotor and Visuotactile Mismatches. *IScience, 23*(3), 100901. https://doi.org/10.1016/j.isci.2020.100901

Salomon, R., Lim, M., Pfeiffer, C., Gassert, R., & Blanke, O. (2013). Full body illusion is associated with widespread skin temperature reduction. *Frontiers in Behavioral Neuroscience, 7*. https://doi.org/10.3389/fnbeh.2013.00065

Serino, A., Alsmith, A., Costantini, M., Mandrigin, A., Tajadura-Jimenez, A., & Lopez, C. (2013). Bodily ownership and self-location: Components of bodily self-consciousness. *Consciousness and Cognition, 22*(4), 1239–1252. https://doi.org/10.1016/j.concog.2013.08.013

Shokur, S., O'Doherty, J. E., Winans, J. A., Bleuler, H., Lebedev, M. A., & Nicolelis, M. A. L. (2013). Expanding the primate body schema in sensorimotor cortex by virtual touches of an avatar. *Proceedings of the National Academy of Sciences, 110*(37), 15121–15126. https://doi.org/10.1073/pnas.1308459110

Slater, M. (2009). Place illusion and plausibility can lead to realistic behaviour in immersive virtual environments. *Philosophical Transactions of the Royal Society B: Biological Sciences, 364*(1535), 3549–3557. https://doi.org/10.1098/rstb.2009.0138

Smith, A. J. T. (2010). Comment: Minimal Conditions for the Simplest Form of Self-Consciousness. In T. Fuchs, H. Sattel, & P. Henningsen (Eds.), The embodied self: Dimensions, coherence, disorders. Schattauer.

Timmermann, C., Roseman, L., Schartner, M., Milliere, R., Williams, L. T. J., Erritzoe, D., Muthukumaraswamy, S., Ashton, M., Bendrioua, A., Kaur, O., Turton, S., Nour, M. M., Day, C. M., Leech, R., Nutt, D. J., & Carhart-Harris,

R. L. (2019). Neural correlates of the DMT experience assessed with multivariate EEG. *Scientific Reports, 9*(1), 1–13. https://doi.org/10.1038/s41598-019-51974-4

Vollenweider, F. X., & Kometer, M. (2010). The neurobiology of psychedelic drugs: Implications for the treatment of mood disorders. *Nature Reviews Neuroscience, 11*(9), 642–651. https://doi.org/10.1038/nrn2884

Wada, M., Takano, K., Ora, H., Ide, M., & Kansaku, K. (2016). The Rubber Tail Illusion as Evidence of Body Ownership in Mice. *Journal of Neuroscience, 36*(43), 11133–11137. https://doi.org/10.1523/JNEUROSCI.3006-15.2016

Walsh, E., Guilmette, D. N., Longo, M. R., Moore, J. W., Oakley, D. A., Halligan, P. W., Mehta, M. A., & Deeley, Q. (2015). Are You Suggesting That's My Hand? The Relation Between Hypnotic Suggestibility and the Rubber Hand Illusion. *Perception, 44*(6), 709–723. https://doi.org/10.1177/0301006615594266

Watts, R., Day, C., Krzanowski, J., Nutt, D., & Carhart-Harris, R. (2017). Patients' Accounts of Increased "Connectedness" and "Acceptance" After Psilocybin for Treatment-Resistant Depression. *Journal of Humanistic Psychology, 57*(5), 520–564. https://doi.org/10.1177/0022167817709585

Weech, S., Kenny, S., & Barnett-Cowan, M. (2019). Presence and Cybersickness in Virtual Reality Are Negatively Related: A Review. *Frontiers in Psychology, 10*. https://doi.org/10.3389/fpsyg.2019.00158

Body Meets Self

An interview with Frédérique de Vignemont

by Raphaël Millière & Carlota Serrahima

Cite as: de Vignemont, F., Millière, R. & Serrahima, C. (2020). Body Meet Self. An interview with Frédérique de Vignemont by Raphaël Millière and Carlota Serrahima. *ALIUS Bulletin*, *4*, 107-133 https://doi.org/10.34700/gza7-yy49

Frédérique de Vignemont
frederique.de.vignemont@ens.fr
Ecole Normale Supérieure,
Paris, France

Raphaël Millière
rm3799@columbia.edu
Columbia University,
New York, USA

Carlota Serrahima
carlota.serrahima@uantwerpen.be
University of Antwerp,
Antwerp, Belgium

Abstract

In this interview, Frédérique de Vignemont discusses her wide-ranging and influential research program on philosophical issues related to bodily awareness. The conversation explores core questions of this research program, such as the existence of a sense of body ownership, the nature of pain and touch, and the role of the peripersonal space, as well as methodological questions regarding the role of empirical evidence in philosophical investigation and the value of arguments from phenomenal contrast in philosophy of mind. In the course of this discussion, Vignemont defends her interpretation of pathological conditions such as somatoparaphrenia, depersonalization disorder, and peripheral deafferentation, as well as experimental bodily illusions such as the rubber hand illusion.

keywords: *bodily awareness*, *sense of body ownership*, *pain*, *touch*, *peripersonal space*, *somatoparaphrenia*, *depersonalization*, *deafferentation*, *rubber hand illusion*, *peripersonal space*

Your research on bodily awareness is one of the most exhaustive and influential in the field of empirically informed philosophy of mind. Bodily awareness is also the topic to which you have devoted the most significant part of your academic production, which covers a very wide range of philosophical issues related to it – from the specificity of particular bodily experiences, such as pain or touch (Vignemont, 2017b), to our capacity for empathy for the bodily sensations of others (Vignemont, 2017a). Within this range of issues, one that is central in your work, and in particular in your recent book *Mind the Body* (Vignemont, 2018a), is the status of bodily awareness as a form of self-consciousness. The book is indeed a deep exploration of the so-called sense of *bodily ownership*, namely of how subjects experience their body

as their own. What motivated you to work on bodily awareness, and on bodily self-consciousness in particular, in the first place?

I wrote my PhD thesis on self-consciousness, and more specifically on immunity to error through misidentification. At the time, bodily self-awareness was just one among other forms of self-awareness and the sense of agency was a lot more popular in the philosophical and the experimental community than the sense of body ownership. Which made it also more interesting. While many were discussing newly found results on action attribution in schizophrenia and delusions of control, there was simply nothing in philosophy on disorders of bodily self-awareness such as somatoparaphrenia (the delusion that one's limb doesn't belong to oneself) or on the rubber hand illusion (the illusion that a rubber hand is one's own hand). Even in cognitive science, these two minimal forms of self-awareness were left disconnected. When I went to work with Professor Patrick Haggard at University College London, my objective was to investigate how agency and ownership interacted. My project, however, ended up working only on the bodily side. The fact is that Queen's Square was a unique place for such a topic, with everyday bringing of new fascinating results. I was lucky to witness and actively participate in the emergence of a new field of research, with a new "body" community that, in fifteen years, was to become so important. For every new exploratory research, there was still very little theory and much conceptual confusion – the perfect challenge for a philosopher…

Philosophers' concerns about self-consciousness have classically covered further aspects of our consciousness of ourselves. One of these aspects is, paradigmatically, whether and how phenomenal awareness of psychological states amounts to self-consciousness proper. For instance, the famous Cartesian dictum states the alleged evidence of the existence of (one's)self on the grounds of states we would now call cognitive, even when in doubt about the existence of the body. Analytic philosophy has of course produced an amazing body of literature addressing this psychological dimension of self-consciousness. Investigations on bodily awareness in contemporary philosophy of mind, however, sometimes run parallel to the insights in this body of literature. In what sense, if any, do you think these two topics, and correspondingly these two areas of research, complement each other?

The body is a material entity located in space and time in the same way as a rock, a tree or a bird. Yet we do not perceive and experience our body like those other objects. What makes it unique is that it bears a special relation to the self, and to self-awareness. However, I want to make it clear that:

(i) all bodily awareness does not have to be self-awareness and
(ii) not all forms of self-awareness have to be embodied.

In my work, I have repeatedly emphasized the fact that one can experience bodily sensations in a part of one's body though one does not experience this body part as being one's own. This is the case in some patients suffering from somatoparaphrenia or from depersonalization. But also, this is most probably the case of some animals that do not have self-referential abilities even at the nonconceptual level. Put it another way, bodily awareness is more widespread in the animal kingdom than self-awareness.

My second claim targets embodied theories of the self, which I do not endorse. The point is not to deny that bodily awareness may be a gateway to self-awareness at the developmental level. But the fact that the body may be at the origin of self-awareness does not entail that every single instance of self-awareness is constitutively embodied. One might reply that in sensory deprivation self-awareness becomes quite thin. This, however, does not suffice to show that self-awareness consists in bodily awareness. What might be necessary for self-awareness is not specifically information about the body, but incoming information in general. The case of sensory deprivation, however, cannot dissociate the two factors. Furthermore, discussions over the notion of immunity to error through misidentification relative to the first-person reveal a clear difference between psychological and bodily self-awareness. Since Evans (Evans 1982), most agree that the self-ascription of bodily judgments is immune to error if grounded in the right way of gaining information about one's body. However, there seems to be a difference between the judgment "I think" and the judgment "my legs are crossed". Immunity directly follows from the peculiarities of mental states but it does not follow from the nature of bodily states. One should not reduce the former to the latter.

> The fact that the body may be at the origin of self-awareness does not entail that every single instance of self-awareness is constitutively embodied.

One of the main distinctive features of your approach is its substantive reliance on empirical data, both from experimental and clinical psychology. Cases such as the rubber hand illusion, somatoparaphrenia, or xenomelia (a disorder characterized by a desire to be disabled or having discomfort with being able-bodied), are now well-known in philosophical discussion of bodily awareness, and this is surely in part because of your own effort to highlight their theoretical import. On the grounds of some of these cases you draw conclusions, for instance, on the ontology of consciousness: on your view, postulating the existence of certain bodily feelings is the best way to explain some features of the empirical cases, against a general principle of phenomenal parsimony (Vignemont, 2018a, section 1.1, 2019a). Your work on the epistemology of bodily awareness appeals to empirical findings as well: they help evaluate the validity of the thesis of the immunity to error through misidentification of bodily self-ascriptions (Vignemont, 2011, 2018a, chap. 3). What would you say is, in general, the methodological value of empirical cases for philosophers? And what would you say is the value of *a priori* philosophical arguments nowadays, in particular for research on bodily awareness?

In my work, I constantly go back and forth between the philosophical and the empirical literature. Cognitive science has grown to such an extent that to neglect what it has to tell us about the mind seems a pity. However, this is not to say that experimental findings are always relevant to philosophical issues. Nor is it to say that one can simply endorse whatever conclusions cognitive neuroscientists reach.

The first step is to decide whether the question that one has in mind can be addressed empirically. For instance, my current research concerns the nature of valence, whether it should be explained in terms of content or in terms of attitude. No experimental result can answer this question. By contrast, I have also been working on spatial awareness. There, the discovery of the specific multisensory and motor properties of peripersonal space in cognitive

“ Experimental findings can be the starting point but they do not replace philosophical arguments. Ideally, they should be the cherry on top, which only empirically confirm what has been found ”

neuroscience is crucial. This discovery shows that we do not process our immediate surroundings in the same way as far space.

The second step for philosophers is to proceed to a task of conceptual clarification: to what extent does peripersonal space differ from egocentric space, from reaching space, or from personal space? Do all studies and all paradigms investigate the same notion? (Vignemont et al., forthcoming) etc. Only once the notion is clearer can one draw the philosophical implications of the findings. For instance, there has been much discussion on enactivism and sensorimotor theories of perception. What new light do these results shed on these theories? How is perception related to action in close space (Vignemont, 2018c)? What is important to keep in mind in philosophy of cognitive science is that at no point should a hypothesis rely exclusively on empirical data. We have recently seen enough that one cannot always replicate results and one should always be very cautious when using them, especially when one is a philosopher with little statistical expertise. Experimental findings can be the starting point but they do not replace philosophical arguments. Ideally, they should be the cherry on top, which only empirically confirm what has been found conceptually.

And the other way around: in your experience of collaboration with neuroscientists and psychologists (e.g., Folegatti et al., 2012; Gouzien et al., 2017), what do you think is the specific contribution that philosophers can make to the cognitive sciences?

I have had the chance to collaborate with many wonderful researchers from various fields, including cognitive psychologists, like Alessandro Farnè and Patrick Haggard, neuroscientists like Tamar Makin and Tania Singer, psychiatrists, like Victor Pitron and Adrienne Gouzien, and even roboticists like Aldo Faisal and Silvestro Micera. Some of our work together was theoretical. For instance, with Tamar Makin and Silvestro Micera, we just

have a short opinion piece accepted in *TICS* on augmentative technology (Makin et al., 2020). We have combined our respective expertise to determine what the most promising path is for the integration of prostheses, arguing for what we call soft embodiment, which we define at the conceptual, neural and computational levels. Other contributions have been more directly experimental. Ideally, some philosophical theories should at least be empirically testable and can be a source of inspiration for cognitive scientists. Furthermore, by constantly anticipating objections, philosophers are well prepared to think of control conditions. Beside direct collaboration, I have found it very interesting to discuss work in progress with junior researchers. As a philosopher, I have more time for reading than experimentalists and this may be helpful to bring the results back into the big picture.

To defend the claim that there is a phenomenology of bodily ownership, you rely in large part on so-called *arguments from phenomenal contrast*. This label was coined by Siegel (2007), although the relevant class of arguments has a long history in philosophy. Arguments from phenomenal contrast are generally used to arbitrate introspective disagreements regarding the existence of a specific kind of phenomenology, by proceeding in two steps:

1. It is argued that a pair of similar experiences E1 and E2 differ – in some small but noticeable way – with respect to their phenomenal character (i.e., what it is like to have them).
2. It is argued that the best explanation of the phenomenal contrast between E1 and E2 is that one experience involves a phenomenal feature F that the other lacks.

In recent years, arguments from phenomenal contrast have been notably used in the debate on the existence of a *sui generis* cognitive phenomenology (e.g., Strawson, 1994; Chudnoff, 2015), and the debate over whether 'high-level' properties - such as the property of being a pine tree - are represented in visual experience (Bayne, 2009; Siegel, 2010).

In your own work, you use arguments from phenomenal contrasts in two complementary ways (Vignemont, 2007, 2013, 2018a, forthcoming). Firstly, you argue that the phenomenal contrast between the experiences produced by the rubber hand illusion experiment in the synchronous (illusory) and asynchronous (non-illusory) conditions, respectively, is best explained by the following hypothesis: subjects feel a phenomenology of ownership over the rubber hand in the synchronous condition, while they lack such

phenomenology in the asynchronous condition. If this hypothesis is true, then the phenomenology of bodily ownership exists *at least in rare experimental conditions*. Secondly, you argue that the phenomenal contrast between a healthy individual's experience of her limb and a patient's experience of her limb in what you call 'disownership syndromes' – which include most prominently somatoparaphrenia, but also depersonalization disorder and other conditions – is best explained by the following hypothesis: in the former (ordinary) condition, but not in the latter (pathological) conditions, one's experience of one's limb involves a phenomenology of bodily ownership. If this hypothesis is true, then the phenomenology of bodily ownership exists not only in rare experimental conditions, but is also prevalent in the ordinary experience of healthy individuals.

Arguments from phenomenal contrast face a well-known challenge: if there is a genuine phenomenal contrast between E1 and E2, it might really be the case that E1 *lacks* a phenomenal feature F that E2 has; but it could also be the case that E1 has an *additional* feature F' that E2 lacks. With respect to the phenomenology of bodily ownership, it could be the case that (a) the experience of the rubber hand in the asynchronous condition of the rubber hand illusion, and (b) the experience of the affected limb in disownership syndromes, do not lack a feature that the contrasted condition has, but instead involve an *additional* and *abnormal* sense of alienation with respect to the relevant limb (Billon & Kriegel, 2015; Chadha, 2018). In fact, there is no shortage of reports from somatoparaphrenic and depersonalized patients that describe a limb as 'alien' or 'strange', or use a lexicon that seems neutral with respect to the two interpretations of the phenomenal contrast (e.g., by saying that the limb is 'dead' or 'rotten'). Do you think this challenge can be addressed, such that the debate over the interpretation of the phenomenal contrast can be settled with a reasonable degree of confidence? If so, what do you think is the most promising kind of evidence that could rule out the alternative interpretation?

First, I should point out that the method of phenomenal contrast is far from perfect. It has come upon heavy criticisms for classic cases (as in the visual experience of pine trees), and it is even more controversial for bodily awareness. As discussed in my latest paper on the phenomenal contrast of ownership, the "all things being equal" rule cannot apply here when it comes to the sense of bodily ownership because the cases we have, such as the rubber hand illusion and disownership syndromes, involve many other differences besides ownership (Vignemont, 2020). That's why, since my *Analysis* paper in

2013, I have always said that a single case taken in isolation is not sufficient to show that there is a phenomenology of ownership and thus to reject the deflationist view (Vignemont, 2013). What we need is to consider a range of cases taken together and abstract what they have in common. But even then, at most what we can offer is an inference to the best explanation. This means that it definitely cannot rule out alternative explanations. What I argue is just that the hypothesis of a phenomenology of ownership can easily account for a range of cases.

Now the hypothesis that by default there is no phenomenology of ownership (hereafter default hypothesis) is a valid alternative interpretation of the evidence, but is it a better one? To settle the debate, we need to know what the arguments are in favour of this view. One could argue that it fits more our introspective reports. As repeatedly emphasized by philosophers from all sides, bodily ownership is not phenomenologically salient under normal condition. The simplest explanation would be that it is because there is no phenomenology of ownership. However, bodily awareness is recessive in general, this is not a problem specific to ownership. Even more generally, what is too familiar always goes to the background of consciousness. And what is more familiar than the fact that this is our own body? There is a real question here but it is an issue about the rules of consciousness in general, and it is not specific to the sense of ownership. Another argument in favour of the default hypothesis is that it seems more parsimonious. But is it really? After all, it involves that in the non-default cases, bodily experiences can represent non-ownership (Chadha, 2018). Advocates of a conservative conception of perception already complain that ownership is a too high-level property for being part of the bodily content, but non-ownership fares even worse. The fact that it is only in some rare cases does not make it more admissible.

As mentioned above, your main argument from phenomenal contrast in favor of the existence of a phenomenology of bodily ownership in the ordinary experience of healthy individuals relies on empirical evidence regarding somatoparaphrenia (SP) (Vignemont, 2007, 2013, 2018a, forthcoming).

SP is a monothematic delusion (typically caused by a brain lesion) characterized by the patients' belief that one of their body parts is not really

theirs. A number of patients also believe that the affected limb belongs to someone else. Your argument from SP starts with the assumption that there is a phenomenal contrast between (a) what it is like for an SP patient to have bodily sensations in the body part whose ownership they deny (e.g., their right hand), and (b) what it is like for a healthy individual to have bodily sensations in the corresponding body part (e.g., their right hand).

SP is associated with a large number of severe motor and somatosensory impairments, which frequently include unilateral neglect, hemiplegia on the contralesional side of the body (paralysis of half of the body, including the affected limb), impairment of the ability to determine the position of one's affected limb through proprioception, hemianaesthesia on the contralesional side (impairment or loss of tactile perception), and hemianopia on the contralesional side (loss of vision in half of the visual field) (Vallar & Ronchi, 2009; Romano & Maravita, 2019).

In many cases, the ability of SP patients to feel bodily sensations at all in their affected limb is heavily impaired. Arguments from phenomenal contrast normally focus on 'minimal pairs' of experiences that are as similar as possible, with the exception of one clear and unique difference. Thus, one might expect the argument from SP to focus on the phenomenal contrast between two experiences involving the same type of bodily sensation (e.g., touch) in the same limb (e.g., the subject's right hand), to bring out a single phenomenal feature that is missing in one experience and present in the other, all else being equal. However, the difference between the bodily experiences of somatoparaphrenic patients and that of healthy individuals are so dramatic that it might be difficult to find such a minimal pair of experiences.

In *Mind the Body*, you acknowledge that the ability to feel touch is frequently affected in SP (2018a, p. 40), but you point out that many SP patients can feel pain in the affected limb, and that in a handful of rare cases these patients can also feel and report touch in the affected limb (e.g., Bottini et al., 2002). You argue that the argument from SP can be set up by using these kinds of cases to bring out the relevant phenomenal contrast. However, there is little evidence that in such cases, the determinate experience of pain (or touch) of an SP patient and the determinate experience of pain (or touch) of a healthy individual differ, if at all, with respect to their phenomenal character. While there is ample evidence that *overall*, SP patients have a range of abnormalities in the way they can experience their affected limb, there is less evidence to support the specific claim that patients with intact ability to feel pain or touch

in the affected limb have a different determinate phenomenology from healthy individuals when they experience pain or touch in the affected limb.

Do you agree with this assessment of available empirical evidence? If so, to what extent do you think the first step of your argument from somatoparaphrenia – establishing the existence of a phenomenal contrast between a minimal pair of bodily experiences – might be affected at all by this assessment?

This is a problem encountered by all attempts to use the phenomenal contrast method: can the contrast be explained by other things? Now in the specific case of ownership, I have collected descriptions not only of somatoparaphrenia, but also of peripheral deafferentation, which is especially interesting, I believe, because it is as pure of a comparison as it can be. Deafferented patients have no brain lesion. This avoids the risk of reasoning deficit, neglect, and other attentional or cognitive disorders, which might be found in brain-lesioned patients or psychiatrist patients. The patients suffered only from a peripheral loss of proprioception and touch. But they can still feel pain and thermal sensations. Now one of them, Ian Waterman, describes how at the beginning, he did not feel his body as being his own (Cole 1995). This, however, did not last. We thus have a contrast between the beginning of the disease and a later stage. One way to describe it is to say that at the early stage, Ian feels pain in *some* legs and that at the late stage, he feels pain in *his* legs. The comparison is not between a patient and healthy subjects, but intra-individual. Now the difference between the two stages is not only a matter of ownership. There is also an agentive contrast. Because of the loss of proprioception, deafferented patients need to learn to exploit vision to replace proprioception to control their bodily movements. So, at the early stage, they have no control over their body and at the late stage, they have regained it. I do not think that the agentive contrast shows that there is no difference at the level of the preserved bodily sensations. Instead, I believe that it can explain the ownership contrast: this temporary loss of agency impacts their preserved bodily sensations, thus explaining the loss of ownership.

To conclude, as I have said again and again, I do not believe that somatoparaphrenia suffices to answer all our questions. What we need to do

is to consider all the cases that are relevant. And it is only taken all together that these various borderline cases can reveal what it is like to experience one's body as one's own.

Another of your arguments from phenomenal contrast in favor of the existence of a phenomenology of bodily ownership relies on evidence provided by the so-called "rubber hand illusion". The rubber hand illusion is a bodily illusion in which a participant sits in front of a fake hand aligned with one's body, while their real hand is hidden behind a screen (Botvinick & Cohen, 1998). Their real hand is subsequently stroked while the fake hand is stroked either synchronously or asynchronously. Thus, participants feel tactile sensations on their real hand while they see the fake hand being stroked at the same time or with some delay. After a few minutes, most participants report the following effects in the synchronous condition but not in the asynchronous condition: (a) the tactile sensation of being stroked feels as if it was located on the rubber hand in front of them rather than on their real hand; and (b) it feels as if the rubber hand was their own hand. These reports typically come in the form of ratings of questionnaire items such as "It seems to me as if the rubber hand were my own hand".

You have argued that bodily experiences elicited by the experimental setup of the rubber hand illusion in the synchronous condition and in the asynchronous condition, respectively, exhibit a phenomenal contrast that lends support to the view that there is a phenomenology of bodily ownership over the fake hand in the former condition. Questionnaire ratings are the primary source of evidence regarding this phenomenal contrast. As you acknowledge yourself, however, the difference in ratings of items related to ownership of the fake hand between asynchronous and synchronous conditions is generally not very impressive: on average, participants are barely in agreement with the idea that the fake hand seemed as if it was theirs during the synchronous condition (Vignemont, 2018a, p. 17; see Longo et al., 2008).

Furthermore, recent work by Peter Lush and colleagues suggests that the rubber hand illusion does not adequately control for demand characteristics – the cues that may convey to participants the experimental outcome or response that the experimenter expects or desires (Lush et al., 2019; Lush, 2020; Roseboom & Lush, 2020) Demand characteristics can not only influence participants' behavior, but also change their experience. Thus, expectancies arising from demand characteristics might cause participants to exert – unknowingly – top-down control of phenomenology, similarly to how they might respond to imaginative suggestions. In fact, Lush and colleagues

found a substantial relationship between trait hypnotisability and both implicit (behavioral) and explicit (verbal) measures of the rubber hand illusion (Lush et al., 2019). In follow-up research, they found that participants' expectancies for "control" and "illusion" statements in synchronous and asynchronous conditions of the rubber hand illusion differ similarly to published illusion reports, implying that standard rubber hand illusion control measures do not effectively control for demand characteristics (Lush 2020). These findings suggest that ownership ratings in the rubber hand illusion may reflect implicit imaginative suggestion effects (in line with Alsmith, 2015).

What do you make of these observations about the rubber hand illusion? Do you think they weaken the evidential strength of reports from the rubber hand illusion for your argument?

When the rubber hand illusion started to be systematically tested, the community thought that we had the magic key to experimentally investigate the sense of body ownership. Since then, after more than 20 years of research and hundreds of versions, I believe that we are all less enthusiastic and we wish we could find other experimental paradigms to test bodily self-awareness. Still, despite their limits and weaknesses, it does not mean that we should just give up on bodily illusions. There are lessons to draw from all the results. We just have to be cautious.

It is also interesting to note that there is a strong parallel with the literature on agency, we are just a few years late. As for agency, the influence of top-down factors is more important than we originally thought. With one of my students, Clément Apelian, we have actually compared the respective impact of hypnosis and sensory manipulation for bodily awareness (Apelian & Vignemont, in preparation). Now the fact that an illusion can be influenced by cognitive factors does not entail that the illusion is constitutively cognitive. It just shows that it can be cognitively penetrated, possibly through imagination. Again, it is important to be vigilant about the significance of empirical results.

> "There is a strong parallel with the literature on agency, we are just a few years late."

According to your *affective account* of the phenomenology of ownership, the phenomenology of bodily ownership is a type of affective phenomenology that bears some resemblance with the feeling of familiarity one may have upon seeing well-known faces, such as those of family members and friends. Nonetheless, the affective feeling of bodily ownership is meant to be more specific than the generic feeling of familiarity. Indeed, unlike the latter, it exclusively tracks one object (one's own body), and it has a positive valence that is motivating for action (i.e., it motivates oneself to protect one's body) (Vignemont, 2018a, p. 192). For this reason, you describe this affective phenomenology as a "narcissistic" feeling, namely an awareness of the special significance that one's body has for oneself. As you also put it, in having a bodily experience "one is aware of bodily boundaries *as having a special significance for the self*" (Vignemont, forthcoming).

This affective account of the phenomenology of bodily ownership is intended to address a dilemma raised by Peacocke (2015), that you summarize as follows:

> "if the protective body map represents one's body *qua* one's own, then it presupposes what it is supposed to explain, but if it does not, then one is left with no explanation of the first-personal character of the sense of bodily ownership" (Vignemont, 2018a, p. 204).

To escape the first horn of this dilemma, you argue that bodily experiences that involve a phenomenology of bodily ownership do not explicitly represent the subject's body *as their own*. Instead, the "narcissistic" feeling of ownership represents the subject's body *as the body that matters (or the body that has special significance)* (Vignemont, forthcoming).

Presumably, the "first-personal character" of the phenomenology of bodily ownership cannot refer to it having *de se* content, since it does not represent the subject's body *as their own*. Accordingly, you argue that the first-personal character of the phenomenology of bodily ownership is guaranteed by the "format" or "structure" of narcissistic feelings, rather than by their content. Indeed, you claim that the brain has a "protective body map" that always and exclusively tracks one's own body; as a result, bodily experiences anchored in the protective body map are automatically tagged with a sort of "self-centred glow" (Vignemont, 2018a, p. 205) derived from the body-tracking function of the protective body map. On your view, this "self-centred glow" is not a matter of the bodily experience having *de se* (e.g., first-personal) content, but is rather like the perspectival structure of visual experience: when one sees a tree in

front of oneself, you suggest that one's visual experience represent the tree as being "in front" (*simpliciter*), rather than "in front *of me*". Similarly, when I feel pressure on my hand, my experience represents pressure "on the body that matters", rather than pressure "on *my own* body" (ibid., p. 205).

One might wonder to what extent this analogy illuminates the nature of the "first-personal character" of the phenomenology of bodily ownership. It is certainly plausible that the sensory experience of very simple organisms, and perhaps even some visual experiences in humans, have merely *de hinc* spatial content without *de se* content – e.g., representing something as being in front of *here*, where *here* refers to the location of the point of origin of the sensory experience's egocentric frame of reference (see Peacocke, 2014, 2017; Schellenberg, 2016). However, many authors agree that ordinary visual experiences represent the locations of environmental landmarks with respect to the subject's own location as such; in other words, when one sees a tree, one's visual experience has a content of the type <There is a tree in front of *me*> (e.g., Cassam, 1997; Noë, 2005; Peacocke, 1998, 2014; Schwenkler, 2014). On this view, visual experiences typically have (nonconceptual) *de se* content. If that wasn't the case, then it might be difficult to see in what sense visual experiences have "first-personal character" at all – any more than a self-driving car's representation of its environment from the data provided by its sensors (cameras or LIDAR system) has "first-personal character", simply because it represents spatial properties of the environment in an egocentric frame of reference. Similarly, if bodily experiences represent one's body as the body that matters *simpliciter*, one might wonder where the "first-personal character" of such experiences comes from. To borrow an example from Martin (1995), what would be the difference between having a bodily experience with the content <There is hurt in *this* ankle>, which surely has no "first-personal character", and a bodily experience with the content <There is hurt in *the ankle that* matters>?

On a related note, *mattering and having a special significance* are normally dyadic predicates: something can only matter or have a special significance *to* or *for someone*. Consequently, one might also wonder whether an experience can represent one's body as the body *that matters* without representing it as the body that matters *to me* (or, indeed, *to someone else*).

What do you make of these concerns regarding the idea that the phenomenology of bodily ownership has "first-personal character" without *de se* content?

To answer Peacocke's dilemma, which is about the body representation, all I need to claim is that the protective body map has no de se content. That's not the same as to say that the experience of ownership has no de se content. I actually do not make this latter claim. What I claim is that the relation to the subject should be understood in terms of personal significance, and not of myness. Actually, in my latest paper on ownership in the *Journal of the APA* (Vignemont, 2020), I argue that ownership experiences do have de se content. More specifically, I defend the view that the subject is part of the truth conditions of the content. Now one can defend this view and still leave implicit the self component. Such a move has been made by Perry (1993) about egocentric experiences, for instance. Egocentric terms actually illustrate how dyadic predicates can be represented as being monadic (Campbell, 1994).

However, there is an alternative worth exploring, namely, that the *de se* nature of the experience of ownership follows from the format of bodily experiences instead of its content. This can take two forms. The first option is that there is a *de se* mode of bodily experiences. This is in line with Recanati's (2012) theory of mental files. The second option is that bodily experiences that can ground ownership judgments have a distinctive affective mental paint. This is in line with the attitude-based approach to affective representations (Mitchell, 2019; Deonna & Teroni, 2012).

To conclude, the main objective of the Bodyguard hypothesis was to develop the idea that the phenomenology of ownership is about personal significance. But a lot remains to be done to explain in detail how to analyse this notion of personal significance, at the level of content or of attitude.

In a recent post on *The Brains Blog* (Vignemont, 2019b), you further elucidate the first horn of the dilemma raised by Peacocke (2015) as follows:

> "One may claim that the conceptual mineness content instantiated by the ownership judgment is grounded in a non-conceptual mineness content at the level of the feeling. But this solution seems to simply beg the question and to leave us with no explanation of the mineness content (conceptual or non-conceptual)."

To understand the pull of this objection, it might be helpful to say something more about the nature of the required explanation, namely – what it is about the existence of a (non-conceptual) representation of a body part as one's own that begs for an explanation? Is what we are after an explanation of the origins of such representation through evolution and individual development, or a mechanistic explanation of why such representations feature in the content of any specific bodily experience, or both? Might the kind of explanation you provide for the existence of a representation of a body part *as having a special significance* not account equally well for the existence of a representation of a body part *as one's own*?

The key question is what the difference is between a creature that can represent only "this leg is bent" (Peacocke's degree zero of self-representation) and a creature that nonconceptually represents "my leg is bent" (degree one of self-representation). To reply that the difference is only that at degree one the creature has a non-conceptual grasp of myness does not bring us very far. It is possible to posit myness as an irreducible primitive phenomenal property but I believe we should do so only when all the other attempts have failed. It seems more interesting to try to understand what new abilities are available to the creatures at degree one independently of myness. Another way to ask the question is what is required for one to be aware of one's body as one's own. I have argued that there are at least three abilities: (i) the spatial ability to individuate the boundary of the body; (ii) the affective ability to ascribe a value to one side of the body boundaries; and (iii) a general self-referential ability. It might then seem that I am also begging the question since I appeal to self-referential abilities in the *explanans*. However, this would not be true. My recipe is <personal significance + self = ownership>. I do not pretend to offer an explanation of self-awareness in general. My objective is more modest: I only offer an account of *bodily* self-awareness, this requires the creature to already have the ability to be self-aware. There is thus no circularity. A prediction of my view is thus that an individual who suffers from a complete disruption of self-referential abilities in other domains would not be able to experience her body as her own. This might be the case in depersonalization disorder, in which their sense of body disownership may result from a more general disruption.

> “I do not pretend to offer an explanation of self-awareness in general. My objective is more modest: I only offer an account of bodily self-awareness.”

In the debate over the existence of a phenomenology of bodily ownership, at least four pairs of labels have been used to distinguish between available positions: (a) "realism" is opposed to "eliminativism" (Gallagher, 2017), (b) "inflationism" is opposed to "deflationism" (Bermúdez, 2011; Vignemont, 2013; Gallagher, 2017, 2019; Serrahima, 2019); (c) "antireductionism" is opposed to "reductionism" (Martin, 1995); and (d) "liberalism" is opposed to "conservatism" (Vignemont, 2018a). In your own work, you have labelled your own view as "liberal" (Vignemont 2018a, p. 13) and "deflationary" (Vignemont, forthcoming). In so far as you defend the existence of a phenomenology of bodily ownership, your position can be adequately described as "realist"; and in so far as you argue that the phenomenology of ownership is not "an irreducible mineness quality" (Vignemont 2018a, p. 48), but is instead reducible to an affective quality, your view can also be presumably described as "reductionist".

How do you understand these various labels, as qualifications of your view of the phenomenology of bodily ownership? Relatedly, do you think that some or all of the four dichotomies listed above are equivalent, or partially overlap, or are orthogonal?

I think that Bermúdez's original distinction between deflationary and inflationary views played an important role by starting a debate at the beginning. However, I've struggled for many years to exactly understand what he meant by it. I originally thought that inflationists claim that there is a phenomenology of ownership while deflationists claim that there is no such thing. But this is not the issue and both sides, as long as they are not eliminativist, agree that it feels like something to be aware of one's body as one's own. Nor was the issue whether one could give a reductionist account of the phenomenology of ownership. Indeed, since the beginning Bermúdez acknowledged that my theory was reductionist and yet I was called an inflationist. The crucial question then is whether one defends the existence of a feeling of myness or not. Indeed, all the objections that Bermúdez offers against inflationism are about myness, and only about that. In brief, inflationists are pro-myness and deflationists are anti-myness. However, in my first paper on ownership (Vignemont, 2007), I did not even use the term of myness (I double-checked), though it was on the basis of this paper that Bermúdez decided that I was an inflationist. More explicitly, in Mind the body (2018a), I clearly criticize the myness hypothesis. On Bermúdez's

taxonomy, I am thus a deflationist. But most people are actually deflationist and I am not convinced that this taxonomy really helps understanding the disagreements between the various views. I find the classic distinction between liberal and conservative conceptions more informative. It brings the debate on ownership within the wider issue of the admissible content of perception. The crucial question is whether one defends the idea that bodily experiences represent only low-level somatosensory properties (pressure, posture, location, temperature, etc.) or whether they can represent high-level properties. These other properties include myness but not only. It can also include agentive properties and affective properties, such as personal significance and value. In brief, liberals are pro rich bodily content whereas conservatives are anti rich bodily content. Within this taxonomy, I am a liberal. So, to conclude, I am a liberal deflationist.

In your view, pain and touch play complementary roles in the development of the sense of bodily ownership in healthy individuals (Vignemont, 2016, 2017b, 2018b). In a nutshell, touch makes an essentially spatial contribution, whereas pain makes an essentially affective contribution. In touch, we perceive non-bodily objects by being in contact with them, which makes the bodily boundaries especially salient. In virtue of this, touch fundamentally contributes to the individuation of the body with respect to other objects – namely those that fall outside the bodily boundary. In turn, pain adds an affective valence to the body that stands at one of the sides of the perceived boundary, contributing a phenomenology of bodily ownership to bodily sensations best described in terms of *care* or *import*, as discussed above. Roughly put, it is only the body that hurts. In your own words, pain "vividly highlights for the subject that what is inside bodily boundaries matters for the self, for its needs, its comfort, and its preservation" (Vignemont, 2017b, p. 475). Protective behavior is then one central manifestation of the sense of bodily ownership.

For a start, is this a fair summary of your ideas on how touch and pain interact to give raise to a sense of bodily ownership?

This is a perfect summary. The skin is a natural boundary that one can be aware of through touch. However, bodily self-awareness cannot be reduced to the spatial awareness of the body. We need more than the boundaries that touch can provide. We need to know which side of the boundaries we are in. And that's where pain, with its affective component, can play a role.

Your description of how touch contributes to bodily ownership follows Mike Martin's (1992, 1993, 1995). Martin argued that all located bodily sensations convey a sense of the boundaries of the body. Arguably, his own analysis of all bodily sensations as involving a sense of what is inside and what is outside the bodily boundaries is an extension, to somatosensation in general, of the model that he endorses for tactile perception. In a paper with Olivier Massin (Vignemont & Massin, 2015), you have called Martin's model of tactile perception the *template model* of tactile perception. However, you argue *against* the template model, and in favour of an alternative *body map* model of tactile perception. Does your account of bodily ownership rely in any way on the specificity of the body map model of touch, vis à vis the template model? Or do you see these two parts of your research as relatively unrelated?

I am greatly indebted to Martin's work on touch and on the sense of ownership, though we do not reach the same conclusions. The template theory and the body map theory actually do not address the same dimensions of touch, the template theory focusing on the exteroceptive content of touch (e.g. how I feel the circular shape of the glass), and the body map theory focusing on its bodily content (e.g. how I feel pressure on my skin).

For the sense of body ownership, what is directly relevant is the fact that the property of pressure is relational: it involves a force exerting on your body, which is independent of you. In a new work with Olivier Massin, we argue that touch gives us a unique sense of reality of what is touched because it presents it as being both mind-independent and causally efficient (Massin & Vignemont, 2020). Touch informs us that the felt object can move, or have an effect, on other entities. When one actively touches an object, one exerts a force on it and one feels not only one's effort but also its resistance to one's effort: "There is no commoner remark than this, that resistance to our muscular effort is the only sense which makes us aware of a reality independent from ourselves" (James, 1890). Only effortful touch presents us

> "Only effortful touch presents us with the contrast between ourselves as striving agents and an independent causally empowered being that resists our effort. Tactile experiences thus give us the boundary between body and world."

with the contrast between ourselves as striving agents and an independent causally empowered being that resists our effort. Tactile experiences thus give us the boundary between body and world, which is required to draw the boundary of the body, a lot more than the visual experiences of the body because only tactile experiences can give this sense of reality of what is not the body.

In your paper "The first person in pain" (Vignemont, 2018b), you write the following:

> "I would like to suggest that provided a pain is felt as one's own, one will react to it normally, even if the body part in which it is felt is not itself felt as one's own. Feeling pain is always of great concern to me, no matter where I feel it."

In this quote, you leave room for the idea that normal reactions to pain, which include protective behaviour, result from the fact that pain be felt as one's own *mental state*, independently of whether or not one has a sense of ownership for the painful body. This idea, however, seems somehow in tension with your explanation of the development of the sense of bodily ownership. Protective behavior, and the affectivity attached to it, seem constitutive of what it is to experience the body as one's own, in the framework of your view. If the very subjectivity of pain experiences – the fact that they are felt as the subject's experiences – suffices to motivate normal protective behavior, does this not undermine the importance of the inside-outside distinction, provided by touch, for the sense of ownership?

Imagine having a terrible pain in the back. You feel the urge to take a painkiller. This can be conceived as a protective behavior. However, the question is what you protect: your body or your psychological life? Let's imagine now that your pain was actually indicative of a kidney infection. By hiding the symptom, you were actually delaying the diagnosis and thus, the cure of the disease. Many have discussed this killing-the-messenger problem and asked whether it is rational or not to take painkillers. But my point here is simply that there are protective behaviours that primarily concern our mental life, and not our body, and that pain, like other negative emotions, can also elicit this type of response, no matter where we locate our pain. I do not think that we should take protective behavior as being a unified category. For instance, one may say that to put money on the side to make sure you

still have something to survive on when you retire is a protective behavior. But it clearly does not involve the protective body map. As noted in the last chapter of *Mind The Body*, one should be extremely careful when analyzing the way people protect or fail to protect themselves. Even in the case of pain, the relation to pain responses is complex and deserves a detailed treatment.

One topic you have worked extensively on is the nature and cognitive underpinnings of vicarious pain. In a very rich paper with Pierre Jacob (Vignemont & Jacob, 2012), you define different ways in which we may feel the pains of other people. You defend that vicarious pain consists in *imagining* being in pain – in particular, doing so through Enactment-imagination (E-imagination): we mentally simulate the psychological state of the other by activating our pain system offline.

On your proposal, there are two variants of vicarious pain, depending on which of the two subsystems actually involved in pain processing are activated when E-imagining pain. On the one hand, the (offline) activation of the *sensory-discriminative subsystem* gives rise to a form of *pain contagion*: a self-centered type of vicarious pain, in which one imagines how it would be to feel pain in the bodily location of one's own body that maps the other's, and anticipates the sensorimotor consequences of such pain. On the other hand, the (offline) activation of the *affective subsystem* gives rise to *empathetic pain*: a global, non-localised bodily feeling, directed to the other's body. In sum, when it comes to vicarious pains, we have (i) pain-like bodily feelings that presumably one identifies as one's own feelings; (ii) caused by pain that affects the bodies of others; (iii) at least in the case of empathetic pain, actually directed to the bodies of others (that is, not clearly centered on what it is like, or what it would be like, to feel a certain pain *on one's own body*); (iv) yet, and again especially in the case of empathetic pain, involving affective reactions directed to the bodies of others.

In your view, how do these features of vicarious pain interact with the role you ascribe to pain in the emergence of the feeling that a body is one's own? Do we have, for the bodies of others, any feeling similar in any relevant respect to those constitutive of the sense of bodily ownership? If so, how do you reconcile this with the fact that, in normal conditions, the sense of bodily ownership tracks only one's own body?

One of the take home messages of our work with Pierre Jacob is that vicarious pain is not the same as standard pain. No matter how much we might say 'I

feel your pain', the fact is that we don't. Part of the difference comes from the content, as you rightly noted. Part of it comes from the fact that it is under the imaginative mode. Thanks to these differences, one does not normally confuse other people's pain with one's own pain. As a consequence, one does not react to one's vicarious pains in the same way as to one's standard pain. Vicarious pains have thus little consequences for the sense of body ownership. Contagious pain remains self-centered, all about one's body. Empathetic pain, on the other hand, is other-centered but it is almost disembodied.

In a recent talk at the École Normale Supérieure, entitled "Keeping the world at distance", you addressed the notion of social distancing, which has become central to our lives in current pandemic times. In the talk you discussed how social distancing reveals the sensorimotor mechanism of peripersonal space, namely the space immediately surrounding our body. You have argued in print that, because anything occupying peripersonal space might soon be in touch with the body, peripersonal perception has evolved into a type of perception distinctively linked to protective action (Vignemont, 2018c). The notion of peripersonal space thus meshes nicely with your ideas on bodily ownership as bodily care: in a way, the space around our bodies is not completely alien, and our sensorimotor system must be sensitive to it just as it is sensitive to events on or within the bodily boundaries.

However, social distancing is not all about keeping the right distance from others in order to promote self-preservation. We also keep apart from the others in order to protect them from ourselves. Do you think that your notions of peripersonal space and self-care have any implications in this direction, now that attention to other-preservation determines our social interactions?

Most of my recent research has indeed been on peripersonal space and we have just edited the first multidisciplinary volume on the topic with Alessandro Farnè, Andrea Serino and Hong-Yu Wong, entitled *The World At Our Fingertips* (Vignemont et al. 2021). Peripersonal space is a fascinating notion that is the crossroad between perception and action and between self and others. It has an immediate relevance with what is happening right now in the world. Now it is a real open question whether social distancing is more about self-preservation or other-preservation. One can easily imagine that altruism increases social distancing, but could it suffice for social distancing? Imagine you could not catch Covid-19 but you could still contaminate others. How likely is it that your distancing behavior will be more reflexive and less

spontaneous than if you could also get contaminated yourself? But maybe I should trust more in human nature.

What do you think will be hot topics in the philosophical and empirical literature on bodily awareness in the coming years?

More and more people have been talking about virtual reality. However, this is just a tool, a very convenient one, but without the right questions, it won't bring us very far. Many also work on interoception. But though there is something quite fascinating about the constant flow of inner signals we receive, we still need to first provide a theory of interoception before trying to assess what impact it has for our cognitive life. To tell me that high interoceptive score is correlated with this or that does not tell me what interoception is. Is interoception a natural kind or just an umbrella term to refer to many different inner signals? What does interoceptive awareness consist in? Is it perceptual or not? Can it ground knowledge? Does it have a distinctive phenomenology? Can one give a representationalist account of interoceptive experiences? What properties would they then represent? As far as I know, no philosopher has yet tried to answer these questions.

Another fascinating topic is augmentative technology. Sci-fi movies show our future selves endowed with high-tech prostheses and exoskeleton. But could we actually exploit a third artificial arm at no cost for our biological arms? Does the brain process artificial devices more like tools or more like body parts? Or could they have their own sensorimotor and phenomenological signature? And how to best design these artificial devices? These are some questions for the future and I'm sure new ones will arise.

References

Alsmith, A. (2015). Mental Activity and the Sense of Ownership. *Review of Philosophy and Psychology*, 6(4), 881–896. https://doi.org/10.1007/s13164014-0208-1

Apelian, C., & de Vignemont, F. (in preparation). *Hypnotic distortions of body representations.*

Bayne, T. (2009). Perception and the Reach of Phenomenal Content. *The Philosophical Quarterly*, *59*(236), 385–404. https://doi.org/10.1111/j.1467-9213.2009.631.x

Bermúdez, J. L. (2011). Bodily Awareness and Self-Consciousness. In S. Gallagher (Ed.), *The Oxford Handbook of the Self*. Oxford University Press.

Billon, A., & Kriegel, U. (2015). Jaspers' Dilemma: The Psychopathological Challenge to Subjectivity Theories of Consciousness. In R. J. Gennaro (Ed.), *Disturbed Consciousness: New Essays on Psychopathology and Theories of Consciousness* (pp. 29–54). MIT Press.

Bottini, G., Bisiach, E., Sterzi, R., & Vallarc, G. (2002). Feeling touches in someone else's hand. *NeuroReport*, *13*(2), 249.

Botvinick, M., & Cohen, J. (1998). Rubber hands 'feel' touch that eyes see. *Nature*, *391*(6669), 756–756. https://doi.org/10.1038/35784

Campbell, J. (1994). *Past, Space, and Self*. MIT Press.

Cassam, Q. (1997). *Self and World*. Oxford University Press.

Chadha, M. (2018). No-Self and the Phenomenology of Ownership. *Australasian Journal of Philosophy*, *96*(1), 14–27. https://doi.org/10.1080/00048402.2017.1307236

Chudnoff, E. (2015). Phenomenal Contrast Arguments for Cognitive Phenomenology. *Philosophy and Phenomenological Research*, *91*(1), 82–104. https://doi.org/10.1111/phpr.12177

Deonna, J., & Teroni, F. (2012). *The Emotions: A Philosophical Introduction*. *Routledge*.

Evans, G. (1982). The Varieties of Reference. Oxford: Oxford University Press.

Folegatti, A., Farnè, A., Salemme, R., & de Vignemont, F. (2012). The Rubber Hand Illusion: Two's a company, but three's a crowd. *Consciousness and Cognition*, *21*(2), 799–812. https://doi.org/10.1016/j.concog.2012.02.008

Gallagher, S. (2017). Deflationary accounts of the sense of ownership. *The Subject's Matter: Self-Consciousness and the Body*. *The MIT Press, Cambridge MA*, 145–162.

Gallagher, S. (2019). The Senses of a Bodily Self. *ProtoSociology*, *36*, 414–433. https://doi.org/10.5840/protosociology20193616

Gouzien, A., de Vignemont, F., Touillet, A., Martinet, N., De Graaf, J., Jarrassé, N., & Roby-Brami, A. (2017). Reachability and the sense of embodiment in amputees using prostheses. *Scientific Reports*, *7*(1), 4999. https://doi.org/10.1038/s41598-017-05094-6

James, W. (1890). *The principles of psychology*. Henry Holt and Company. http://content.apa.org/books/11059-000

Longo, M. R., Schüür, F., Kammers, M. P. M., Tsakiris, M., & Haggard, P. (2008). What is embodiment? A psychometric approach. *Cognition, 107*(3), 978–998. https://doi.org/10.1016/j.cognition.2007.12.004

Lush, P. (2020). Demand Characteristics Confound the Rubber Hand Illusion. *Collabra: Psychology, 6*(1), 22. https://doi.org/10.1525/collabra.325

Lush, P., Botan, V., Scott, R. B., Seth, A., Ward, J., & Dienes, Z. (2019). *Phenomenological control: Response to imaginative suggestion predicts measures of mirror touch synaesthesia, vicarious pain and the rubber hand illusion*. https://doi.org/10.31234/osf.io/82jav

Makin, T. R., de Vignemont, F., & Micera, S. (2020). Soft Embodiment for Engineering Artificial Limbs. *Trends in Cognitive Sciences, 24*(12), 965–968. https://doi.org/10.1016/j.tics.2020.09.008

Martin, M. G. F. (1992). Sight and Touch. In T. Crane (Ed.), *The Contents of Experience* (pp. 199–201). Cambridge University Press.

Martin, M. G. F. (1993). Sense modalities and spatial properties. In N. Eilan, R. McCarty, & B. Brewer (Eds.), *Spatial representation* (pp. 206–218). Oxford University Press.

Martin, M. G. F. (1995). Bodily Awareness: A Sense of Ownership. In J. L. Bermudez, A. J. Marcel, & N. M. Eilan (Eds.), *The Body and the Self* (pp. 267–289). MIT Press.

Massin, O., & de Vignemont, F. (2020). Unless I Put My Hand into His Side, I Will Not Believe. In D. E. Gatzia & B. Brogaard (Eds.), *The Epistemology of Non-Visual Perception* (p. 165). Oxford University Press.

Mitchell, J. (2019). Affective Representation and Affective Attitudes. *Synthese*, 1–28. https://doi.org/10.1007/s11229-019-02294-7

Noë, A. (2005). *Action in Perception*. MIT Press.

Peacocke, C. (1998). *Being Known*. Oxford University Press.

Peacocke, C. (2014). *The Mirror of the World: Subjects, Consciousness, and Self-Consciousness*. Oxford University Press.

Peacocke, C. (2015). Perception and the First Person. In M. Matthen (Ed.), *The Oxford Handbook of Philosophy of Perception*. http://www.oxfordhandbooks.com/view/10.1093/oxfordhb/9780199600472.001.0001/oxfordhb-9780199600472-e-022

Peacocke, C. (2017). Philosophical Reflections on the First Person, the Body, and Agency. In F. de Vignemont & A. Alsmith (Eds.), *The Subject's Matter: Self-Consciousness and the Body* (pp. 289–310). MIT Press.

Perry, J. (1993). *The Problem of the Essential Indexical: And Other Essays*. Oxford University Press.

Recanati, F. (2012). *Mental Files*. Oxford University Press.

Romano, D., & Maravita, A. (2019). The dynamic nature of the sense of ownership after brain injury. Clues from asomatognosia and somatoparaphrenia. *Neuropsychologia, 132*, 107119. https://doi.org/10.1016/j.neuropsychologia.2019.107119

Roseboom, W., & Lush, P. (2020). *Serious problems with interpreting rubber hand illusion experiments*. https://doi.org/10.31234/osf.io/uhdzs

Schellenberg, S. (2016). De Se Content and De Hinc Content. *Analysis, 76*(3), 334–345. https://doi.org/10.1093/analys/anw019

Schwenkler, J. (2014). Vision, Self-Location, and the Phenomenology of the 'Point of View.' *Noûs, 48*(1), 137–155. https://doi.org/10.1111/j.1468-0068.2012.00871.x

Serrahima, C. (2019). *My Body is the Subject's Body. In Defence of Experientialism about the Sense of Bodily Ownership* [University of Barcelona]. http://diposit.ub.edu/dspace/handle/2445/137017

Siegel, S. (2007). How Can We Discover the Contents of Experience? The Southern Journal of Philosophy, 45(S1), 127–142. https://doi.org/10.1111/j.2041-6962.2007.tb00118.x

Siegel, S. (2010). *The Contents of Visual Experience*. Oxford University Press USA.

Strawson, G. (1994). *Mental Reality*. MIT Press.

Vallar, G., & Ronchi, R. (2009). Somatoparaphrenia: A body delusion. A review of the neuropsychological literature. *Experimental Brain Research, 192*(3), 533–551. https://doi.org/10.1007/s00221-008-1562-y

de Vignemont, F. (forthcoming). The phenomenology of bodily ownership. In M. Garcia-Carpintero & M. Guillot (Eds.), *The sense of mineness*. Oxford University Press.

de Vignemont, F. (2007). Habeas Corpus: The Sense of Ownership of One's Own Body. *Mind & Language, 22*(4), 427–449. https://doi.org/10.1111/j.1468-0017.2007.00315.x

de Vignemont, F. (2011). Bodily immunity to error. In S. Prosser & F. Recanati (Eds.), *Immunity to error through misidentification: New essays* (pp. 1–27). Cambridge University Press.

de Vignemont, F. (2013). The mark of bodily ownership. *Analysis, 73*(4), 643–651. https://doi.org/10.1093/analys/ant080

de Vignemont, F. (2016). Pain and the Spatial Boundaries of the Bodily Self. In L. Garcia-Larrea (Ed.), *Pain and the Conscious Brain*. IASP.

de Vignemont, F. (2017a). Can I see your pain? An evaluative model of pain perception. In J. Corns (Ed.), *Can I see your pain? An evaluative model of pain perception* (pp. 255–265). Routledge.

de Vignemont, F. (2017b). Pain and Touch. *The Monist, 100*(4), 465–477. https://doi.org/10.1093/monist/onx022

de Vignemont, F. (2018a). *Mind the Body: An Exploration of Bodily Self-Awareness*. Oxford University Press.

de Vignemont, F. (2018b). The First-Person in Pain. In M. Brady, D. Bain, & J. Corns (Eds.), *Philosophy of Pain. Unpleasantness, Emotion, and Deviance*. Routledge.

de Vignemont, F. (2018c). Peripersonal perception in action. *Synthese*. https://doi.org/10.1007/s11229-018-01962-4

de Vignemont, F. (2019a). Against Phenomenal Parsimony: A Plea for Bodily Feelings. In A. I. Goldman & B. P. McLaughlin (Eds.), *Metaphysics and Cognitive Science* (pp. 268–283). Oxford University Press.

de Vignemont, F. (2019b). Mind the body (4) What kind of first-personal content? In *The Brains Blog*. http://philosophyofbrains.com/2019/02/28/mind-the-body-4-what-kind-of-first-personal-content.aspx

de Vignemont, F. (2020). What Phenomenal Contrast for Bodily Ownership? *Journal of the American Philosophical Association, 6*(1), 117–137. https://doi.org/10.1017/apa.2019.34

de Vignemont, F., & Jacob, P. (2012). What Is It like to Feel Another's Pain? *Philosophy of Science, 79*(2), 295–316. https://doi.org/10.1086/664742

de Vignemont, F., & Massin, O. (2015). Touch. In M. Matthen (Ed.), *The Oxford Handbook of Philosophy of Perception* (pp. 294–313). Oxford University Press.

de Vignemont, F., Serino, A., Wong, H. Y., & Farnè, A. (forthcoming). Peripersonal space: A special way of representing space. *In The world at our fingertips: A multidisciplinary investigation of peripersonal space*. Oxford University Press. ISBN: 9780198851738

ALIUS BULLETIN

exploring the diversity of consciousness

n°4

DECEMBER 2020

aliusresearch.org