Frontiers | Research Topics

Rising stars in: Consciousness research 2021

Edited by
Antonino Raffone, Narayanan Srinivasan, Luca Simione,
Xerxes D. Arsiwalla and Johannes Kleiner

Published in
Frontiers in Psychology

ISSN 1664-8714
ISBN 978-2-8325-3851-7
DOI 10.3389/978-2-8325-3851-7

About Frontiers

Frontiers is more than just an open access publisher of scholarly articles: it is a pioneering approach to the world of academia, radically improving the way scholarly research is managed. The grand vision of Frontiers is a world where all people have an equal opportunity to seek, share and generate knowledge. Frontiers provides immediate and permanent online open access to all its publications, but this alone is not enough to realize our grand goals.

Frontiers journal series

The Frontiers journal series is a multi-tier and interdisciplinary set of open-access, online journals, promising a paradigm shift from the current review, selection and dissemination processes in academic publishing. All Frontiers journals are driven by researchers for researchers; therefore, they constitute a service to the scholarly community. At the same time, the *Frontiers journal series* operates on a revolutionary invention, the tiered publishing system, initially addressing specific communities of scholars, and gradually climbing up to broader public understanding, thus serving the interests of the lay society, too.

Dedication to quality

Each Frontiers article is a landmark of the highest quality, thanks to genuinely collaborative interactions between authors and review editors, who include some of the world's best academicians. Research must be certified by peers before entering a stream of knowledge that may eventually reach the public - and shape society; therefore, Frontiers only applies the most rigorous and unbiased reviews. Frontiers revolutionizes research publishing by freely delivering the most outstanding research, evaluated with no bias from both the academic and social point of view. By applying the most advanced information technologies, Frontiers is catapulting scholarly publishing into a new generation.

What are Frontiers Research Topics?

Frontiers Research Topics are very popular trademarks of the *Frontiers journals series*: they are collections of at least ten articles, all centered on a particular subject. With their unique mix of varied contributions from Original Research to Review Articles, Frontiers Research Topics unify the most influential researchers, the latest key findings and historical advances in a hot research area.

Find out more on how to host your own Frontiers Research Topic or contribute to one as an author by contacting the Frontiers editorial office: frontiersin.org/about/contact

Rising stars in: Consciousness research 2021

Topic editors
Antonino Raffone — Sapienza University of Rome, Italy
Narayanan Srinivasan — Indian Institute of Technology Kanpur, India
Luca Simione — UNINT - Università degli studi Internazionali di Roma, Italy
Xerxes D. Arsiwalla — Pompeu Fabra University, Spain
Johannes Kleiner — Ludwig Maximilian University of Munich, Germany

Citation
Raffone, A., Srinivasan, N., Simione, L., Arsiwalla, X. D., Kleiner, J., eds. (2023). *Rising stars in: Consciousness research 2021*. Lausanne: Frontiers Media SA.
doi: 10.3389/978-2-8325-3851-7

Table of contents

frontiers | Frontiers in Psychology

TYPE Editorial
PUBLISHED 16 May 2023
DOI 10.3389/fpsyg.2023.1205982

OPEN ACCESS

EDITED AND REVIEWED BY
Adam B. Barrett,
University of Sussex, United Kingdom

*CORRESPONDENCE
Xerxes D. Arsiwalla
x.d.arsiwalla@gmail.com

RECEIVED 14 April 2023
ACCEPTED 25 April 2023
PUBLISHED 16 May 2023

CITATION
Arsiwalla XD, Srinivasan N, Simione L, Kleiner J and Raffone A (2023) Editorial: Rising stars in: consciousness research 2021.
Front. Psychol. 14:1205982.
doi: 10.3389/fpsyg.2023.1205982

Editorial: Rising stars in: consciousness research 2021

Xerxes D. Arsiwalla[1,2]*, Narayanan Srinivasan[3], Luca Simione[4,5], Johannes Kleiner[2,6,7] and Antonino Raffone[8]

[1]Department of Information and Communication Technologies, Pompeu Fabra University, Barcelona, Spain, [2]Association for Mathematical Consciousness Science, Munich, Germany, [3]Department of Cognitive Science, Indian Institute of Technology Kanpur, Kanpur, India, [4]Institute of Cognitive Sciences and Technologies, Consiglio Nazionale delle Ricerche, Rome, Italy, [5]Faculty of Interpreting and Translation, UNINT Università degli Studi Internazionali di Roma, Rome, Italy, [6]Munich Center for Mathematical Philosophy, Ludwig Maximilian University of Munich, Munich, Germany, [7]Munich Graduate School of Systemic Neurosciences, Ludwig Maximilian University of Munich, Munich, Germany, [8]Department of Psychology, Sapienza University of Rome, Rome, Italy

KEYWORDS

contents of consciousness, sense of self, embodied cognition, hypnosis, meditation, criticality, higher-order theories, monism

Editorial on the Research Topic
Rising stars in: consciousness research 2021

The Rising Stars topic serves as a forum for new directions and emerging trends in consciousness research, showcasing high-quality work of early-stage researchers. Compared to platforms highlighting dominant paradigms and established viewpoints (Srinivasan et al., 2023), this Research Topic presents new independent ideas with promising scientific impact. It highlights the diversity of research performed across the entire breadth of consciousness science, including advances in theory, experiment, methodology and clinical practice.

The article by Oblak et al. report an exploratory phenomenological study on the experience of individuals during task performance, as opposed to purely behavioral assessments (reaction-times/accuracy). The authors used a working memory visuo-spatial change detection task to explore experiences elicited by and accompanying task performance. Implementing contemporary phenomenological and qualitative methodology, they gathered subjective reports during the memory task. Qualitative reports revealed rich experiential landscapes associated with task-performance, suggesting a distinction between two broad classes of experience: phenomena at the front of consciousness and background feelings. The former captured cognitive strategies and aspects of metacognition, whereas the latter encapsulated difficult-to-detect aspects of experience that comprise the overall sense of experience (bodily feelings, emotional atmosphere, mood). This study focused primarily on background feelings of subjects and the nature of the elicited experience.

Bréchet's article looks at how personal memories and bodily-cues influence our sense of self. Bréchet explores the underlying characteristics of self-consciousness and its relation to bodily signals and episodic memory. This article outlines recent behavioral and neuroimaging evidence indicating that bodily cues play a fundamental role in autobiographical memory. The author also discusses emerging concepts regarding the current understanding of bodily-self and autobiographical-self, and their links to self-consciousness. These ideas can further help bridge theoretical notions of self (Metzinger, 2004) to empirical underpinnings.

The article by Dapor et al. investigates bodily states in relation to consciousness. They reviewed the boundaries between the physical and social world, and shed insights on the intersection between those. This review focused on interactions between sensory inputs and social cognition in visual perception mediated by body indices, such as gait and posture. For instance, depending on our arm's length, height and capacity of movement, we create our own image of the world based on a continuous compromise between sensory inputs and expected behavior. The authors propose that we use our bodies as natural "rulers" to measure the physical and social world around us. This goes beyond approaches that define perception as stimulus-centered, and emphasizes an embodied agent-based perspective, which closely resonates with paradigms of enactivism and predictive coding (Clark, 2013; Varela et al., 2016).

Toward improving the quality of life of post-treatment cancer patients, Grégoire et al. presented a study protocol for administering complementary therapeutic approaches in addition to standard cancer therapies. These included hypnosis, meditation and self-induced cognitive trance. The authors detail the protocol of a preference-based longitudinal controlled trial that assessed the effectiveness of the stated interventions on the common symptom cluster associated with oncological patients: cancer-related fatigue, emotional distress, sleep, pain, and cognitive difficulties using questionnaires and neurobiological measures.

Schmidt's article shows how hypnosis elicits positive feelings of safety for people undergoing substance abuse therapy or those in highly stressful environments, such as intensive care units. Patients tolerated non-invasive ventilation much better when they got the suggestion to feel safe. The effects of positive therapeutic suggestions delivered during hypnosis even persist afterwards. Post-hypnotic suggestions are associations between a certain emotional state and a trigger that elicits this emotional state after hypnosis is over. Post-hypnotic suggestions of safety were shown to be effective weeks after therapy. Such studies underscore the potential of manipulating non-standard states of consciousness for realizing new clinical therapies applicable to a broad range of diseases.

The article by Rabuffo et al. propose spontaneous neuronal avalanches as correlates of access consciousness. This draws upon the idea that the brain self-regulates around a critical point, at the edge of a phase transition (Cocchi et al., 2017). This has also been linked to optimal information processing and consciousness (Werner, 2009; Shew et al., 2011; Arsiwalla and Verschure, 2016; Arsiwalla et al., 2017). Based on the observation that spontaneous brain activity dynamically switches between epochs of segregation and large-scale integration of information, the authors hypothesize a brain-state dependence of conscious access, whereby the presence of either segregated or integrated states marks distinct modes of information processing. They propose a test experiment to validate the hypothesis that conscious access occurs in aperiodic cycles; namely, alternating windows where new incoming information is collected but not experienced, to punctuated short-lived integration events, where conscious access to previously collected content occurs. In particular, they suggest that integration events correspond to neuronal avalanches, which are collective bursts of neuronal activity ubiquitously observed in electrophysiological recordings. If confirmed, this framework could link the physics of spontaneous cortical dynamics to the concept of ignition (Mashour et al., 2020) within the global neuronal workspace theory, whereby conscious access correlates to bursts of neuronal activity.

Kirkeby-Hinrup's article discusses the plausibility of misrepresentation of conscious states in the context of higher-order theories. The argument relies on the assumption that conscious states are generated by processes in the brain. The underlying idea is that if the brain generates conscious states then misrepresentations can occur. The reason for this is that brain states can be corrupted and, accordingly, a conscious state that is at least partly caused by a corrupted brain state may be a misrepresentation. Furthermore, given that corruption of neural states is both possible and relatively frequent, the author argues that it is plausible that occasionally such corruption may result in misrepresentation.

The article by Lahav and Neemeh proposes that phenomenal consciousness manifests as a relativistic, rather than an absolute property. These authors contend that both the dualist and illusionist positions are flawed because those tacitly assume consciousness to be an absolute property that is observer-independent. This work proposes a conceptual and a mathematical argument for a relativistic theory of consciousness in which a system either has or doesn't have phenomenal consciousness with respect to some observer. The authors argue that in the frame of reference of the cognitive system, phenomenal properties will be observable (first-person perspective) and in other frames of reference it will not (third-person perspective). These two cognitive frames of reference are both deemed correct, just as the case of an observer claiming to be stationary when another would conclude that the former has constant velocity. Assuming consciousness is a relativistic phenomenon, neither observer's position is privileged. They both describe the same reality. This partially relates to reflexive monism (Velmans, 2009). The authors then seek to bridge the explanatory gap and address the hard problem.

In summary, this Research Topic has brought together diverse works addressing important issues in consciousness research such as the phenomenology of experience, the empirical underpinnings of the self, the role of sociality, the use of non-standard states of consciousness for designing new clinical therapies, a possible mechanism bridging cortical dynamics to access consciousness, an argument on misrepresentation in the context of higher-order theories, and a new hypothesis advocating observer-dependence of phenomenological properties of consciousness. Given the broad scope and rapid growth of the science of consciousness, it is imperative that independent perspectives different from conventional paradigms, are provided with a suitable platform for dissemination. This Research Topic addresses that need.

Author contributions

All authors listed have made a substantial, direct, and intellectual contribution to the work and approved it for publication.

Acknowledgments

We acknowledge the wholehearted help of the Frontiers in Psychology team and all the authors who contributed to this Research Topic.

Conflict of interest

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

Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

References

Arsiwalla, X. D., Mediano, P. A., and Verschure, P. F. (2017). Spectral modes of network dynamics reveal increased informational complexity near criticality. *Proc. Comput. Sci.* 108, 119–128. doi: 10.1016/j.procs.2017.05.241

Arsiwalla, X. D., and Verschure, P. (2016). "High integrated information in complex networks near criticality," in *Artificial Neural Networks and Machine Learning - ICANN 2016. Lecture Notes in Computer Science, vol 9886*, eds A. Villa, P. Masulli, and A. Pons Rivero (Cham: Springer). doi: 10.1007/978-3-319-44778-0_22

Clark, A. (2013). Whatever next? Predictive brains, situated agents, and the future of cognitive science. *Behav. Brain Sci.* 36, 181–204. doi: 10.1017/S0140525X12000477

Cocchi, L., Gollo, L. L., Zalesky, A., and Breakspear, M. (2017). Criticality in the brain: a synthesis of neurobiology, models and cognition. *Progr. Neurobiol.* 158, 132–152. doi: 10.1016/j.pneurobio.2017.07.002

Mashour, G. A., Roelfsema, P., Changeux, J.-P., and Dehaene, S. (2020). Conscious processing and the global neuronal workspace hypothesis. *Neuron* 105, 776–798. doi: 10.1016/j.neuron.2020.01.026

Metzinger, T. (2004). *Being No One: The Self-Model Theory of Subjectivity*. Cambridge, MA: MIT Press. doi: 10.7551/mitpress/1551.001.0001

Shew, W. L., Yang, H., Yu, S., Roy, R., and Plenz, D. (2011). Information capacity and transmission are maximized in balanced cortical networks with neuronal avalanches. *J. Neurosci.* 31, 55–63. doi: 10.1523/JNEUROSCI.4637-10.2011

Srinivasan, N., Simione, L., Arsiwalla, X. D., Kleiner, J., and Raffone, A. (2023). Editorial: insights in consciousness research 2021. *Front. Psychol.* 14, 1182690. doi: 10.3389/fpsyg.2023.1182690

Varela, F. J., Thompson, E., and Rosch, E. (2016). *The Embodied Mind: Cognitive Science and Human Experience (Revised Edition)*. Cambridge, MA: MIT Press. doi: 10.7551/mitpress/9780262529365.001.0001

Velmans, M. (2009). *Understanding Consciousness*. London: Routledge. doi: 10.4324/9780203882726

Werner, G. (2009). Viewing brain processes as critical state transitions across levels of organization: neural events in cognition and consciousness, and general principles. *Biosystems* 96, 114–119. doi: 10.1016/j.biosystems.2008.11.011

frontiers in Psychology

STUDY PROTOCOL
published: 10 February 2022
doi: 10.3389/fpsyg.2022.807741

Hypnosis, Meditation, and Self-Induced Cognitive Trance to Improve Post-treatment Oncological Patients' Quality of Life: Study Protocol

Charlotte Grégoire[1,2]*, Nolwenn Marie[1,2], Corine Sombrun[3], Marie-Elisabeth Faymonville[1,4], Ilios Kotsou[5], Valérie van Nitsen[5], Sybille de Ribaucourt[5], Guy Jerusalem[6], Steven Laureys[2,7], Audrey Vanhaudenhuyse[1,8†] and Olivia Gosseries[1,2,7†]

[1]Sensation and Perception Research Group, GIGA Consciousness, University of Liège, Liège, Belgium, [2]Coma Science Group, GIGA Consciousness, University of Liège, Liège, Belgium, [3]Trance Science Research Institute, Paris, France, [4]Arsène Burny Cancerology Institute, CHU Liège, Liège, Belgium, [5]Faculty of Psychology and Educational Sciences, Free University of Brussels and Emergences Association, Brussels, Belgium, [6]Medical Oncology Department, CHU Liège and University of Liège, Liège, Belgium, [7]Centre du Cerveau[2], CHU Liège, Liège, Belgium, [8]Algology Interdisciplinary Center, CHU Liège, Liège, Belgium

OPEN ACCESS

Edited by:
Gregor Weissflog, Leipzig University, Germany

Reviewed by:
Susan Koranyi, University Hospital Leipzig, Germany
Anja Mehnert-Theuerkauf, University Hospital Leipzig, Germany

****Correspondence:***
Charlotte Grégoire
ch.gregoire@uliege.be
orcid.org/0000-0003-1837-6260

†These authors have contributed equally to this work and share senior authorship

Specialty section:
This article was submitted to Psycho-Oncology, a section of the journal Frontiers in Psychology

Received: *02 November 2021*
Accepted: *05 January 2022*
Published: *10 February 2022*

Citation:
Grégoire C, Marie N, Sombrun C, Faymonville M-E, Kotsou I, van Nitsen V, de Ribaucourt S, Jerusalem G, Laureys S, Vanhaudenhuyse A and Gosseries O (2022) Hypnosis, Meditation, and Self-Induced Cognitive Trance to Improve Post-treatment Oncological Patients' Quality of Life: Study Protocol. Front. Psychol. 13:807741. doi: 10.3389/fpsyg.2022.807741

Introduction: A symptom cluster is very common among oncological patients: cancer-related fatigue (CRF), emotional distress, sleep difficulties, pain, and cognitive difficulties. Clinical applications of interventions based on non-ordinary states of consciousness, mostly hypnosis and meditation, are starting to be investigated in oncology settings. They revealed encouraging results in terms of improvements of these symptoms. However, these studies often focused on breast cancer patients, with methodological limitations (e.g., small sample size, no control group, and no follow-up). Another non-ordinary state of consciousness may also have therapeutic applications in oncology: self-induced cognitive trance (SICT). It seems to differ from hypnosis and meditation, as it involves the body more directly. Thus, investigating its clinical applications, along with hypnosis and meditation interventions, could improve available therapeutic options in oncology. This article details the study protocol of a preference-based longitudinal controlled superiority trial aiming to assess the effectiveness of 3 group interventions (hypnosis, meditation, and SICT) to improve oncological patients' quality of life, and more specifically CRF, emotional distress, sleep, pain, and cognitive difficulties (primary outcomes).

Methods and analysis: A power analysis required a total sample of 160 patients. Main inclusion criteria are: cancer diagnosis, active treatments completed for less than a year, no practice of hypnosis, meditation, or SICT, and presence of at least one of these four symptoms: fatigue, sleep difficulties, depression, or anxiety. Each participant will choose the intervention in which they want to participate (hypnosis, mindful self-compassion meditation, SICT, or no intervention—control group). To test the effectiveness of the interventions, data will be collected by questionnaires and neurobiological measures and directly from the medical record at four time points: before inclusion in the study (baseline); immediately after the intervention; and at 3- and 12-month follow-up. The longitudinal data in each group will then be measured.

Discussion: In addition to standard cancer therapies, there is a growing interest from patients in complementary approaches, such as hypnosis, meditation, and SICT. The results of this study will be useful to increase knowledge about short- and long-term effectiveness of 3 group interventions for CRF, emotional distress, sleep, pain, and cognitive difficulties in patients with different cancers.

Clinical Trial Registration: ClinicalTrials.gov/ (NCT04873661). Retrospectively registered on the 29th of April 2021. url: https://clinicaltrials.gov/ct2/show/NCT04873661

Keywords: oncology, group intervention, hypnosis, meditation, self-induced cognitive trance, neurophysiology, neurophenomenology

INTRODUCTION

The presence of one symptom cluster is particularly well-documented among patients who suffer from cancer (Dodd et al., 2010; Die Trill, 2013; Loh et al., 2018): cancer-related fatigue (CRF), sleep difficulties, emotional distress, and pain. Cognitive impairments are also frequently associated with this symptom cluster (Pullens et al., 2010; Joly et al., 2011; Sanford et al., 2014). A meta-analysis revealed that 52% of patients report CRF, this proportion ranging from 14 to 100% according to the studies (Ma et al., 2020). CRF can be defined as "a distressing persistent, subjective sense of physical, emotional, and/or cognitive tiredness related to cancer or cancer treatment that is not proportional to recent activity and interferes with usual functioning" (Koh et al., 2019). CRF has a lot of social, financial, and functional consequences (Jones et al., 2016; Ma et al., 2020). Emotional distress, endured by a large proportion of patients with cancer as well (Mehnert et al., 2018; Götze et al., 2020), can be defined as "a multifactorial, unpleasant experience of a psychologic (i.e., cognitive, behavioral, and emotional), social, spiritual, and/or physical nature that may interfere with the ability to cope effectively with cancer, its physical symptoms, and its treatment" (Mitchell et al., 2011; Riba et al., 2019). It negatively influences treatment adherence (Lin et al., 2017) and results (Batty et al., 2017), as well as patient's general quality of life (Achimas-Cadariu et al., 2015; Riba et al., 2019). Concerning sleep difficulties, their prevalence among patients with cancer during or after their treatment varies between 32 and 61% (Santoso et al., 2019; Schieber et al., 2019; Hoang et al., 2020; Gonzalez et al., 2021) and is approximately three times higher than in the general population (Kwak et al., 2020). Cancer-related pain is reported by more than 50% of patients, which severely impacts quality of life, adherence to treatment, satisfaction with care, and survival (Neufeld et al., 2017). Finally, cognitive difficulties may be linked to the disease itself, the treatments received or the patient characteristics, and their prevalence can be up to 75% in some kinds of cancers (Joly et al., 2011).

All these difficulties can last for years after treatment completion and, despite their prevalence and severe impact, they remain underdiagnosed and undertreated (Joly et al., 2011; Die Trill, 2013; de Vries and Stiefel, 2014; Riba et al., 2019). In oncology settings, there is a growing interest in complementary approaches, such as "mind–body" interventions, to relieve these symptoms in a non-pharmacological way. These interventions aim to impact two interconnected levels: psychological and physical (Sawni and Breuner, 2017), and they allow patients to take back the control over their health and to regain hope (Eliott et al., 2008). Indeed, complementary and alternative medicine is estimated to be used by 43% of patients with cancer (Horneber et al., 2012; John et al., 2016), and it is starting to be studied scientifically, for instance with trials investigating the benefits of hypnosis and meditation (Spiegel, 1985; Kim et al., 2013; Cramer et al., 2015; Carlson et al., 2017; Grégoire et al., 2017, 2020a; Arring et al., 2019; Ngamkham et al., 2019; Suh et al., 2021).

Hypnosis

Hypnosis can be defined as "a state of consciousness involving focused attention and reduced peripheral awareness characterized by an enhanced capacity for response to suggestion" (Elkins et al., 2015). It is characterized by four main components: absorption (in an imaginative experience), dissociation (from the environment), suggestibility (to the suggestions made by the therapist; Spiegel, 1991), and automaticity (non-voluntary response relevant to the content of a communication intended to be a suggestion; Weitzenhoffer, 2002). Hypnosis modulates self-awareness and decreases environmental awareness (Demertzi et al., 2011, 2015). At a neurophysiological level, it is associated with increased power in the theta band and changes in gamma activity (direction inconsistent among studies), which play a critical role in the emotional limbic circuits (Vanhaudenhuyse et al., 2020). Hypnosis also modifies the functional connectivity in the default mode network (linked to self-awareness) and in the external network (linked to the environmental awareness; Vanhaudenhuyse et al., 2014), as well as the structural connectivity

Abbreviations: ANOVA, Analysis of variance; BFI, Big Five Inventory; CERQ, Cognitive Emotion Regulation Questionnaire; CEQ, Creative Experiences Questionnaire; CHU, "Centre hospitalier universitaire" (University Hospital); CRF, Cancer-related fatigue; ECG, Electrocardiogram; EEG, Electroencephalogram; EMG, Electromyogram; FACT-Cog, Functional Assessment of Cancer Therapy - Cognitive Function; GDPR, General Data Protection Regulation; HADS, Hospital Anxiety and Depression Scale; Hrs, hours; ISI, Insomnia Severity Index; MAC, Mental adjustment to cancer; MCQ, Memory Characteristics Questionnaire; MEQ, Mystical Experience Questionnaire; MFI, Multidimensional fatigue inventory; MPFI, Multidimensional Psychological Flexibility Index; MSC, Mindful self-compassion; SICT, Self-induced cognitive trance; VAS, Visual analogue scale.

between brain areas, and the hemispheric asymmetry (Landry et al., 2017). It also impacts cortical and subcortical areas involved in pain modulation (Vanhaudenhuyse et al., 2014; De Benedittis, 2021; Bicego et al., 2021c).

To summarize empirical findings about hypnosis mechanisms, Jensen et al. proposed a biopsychosocial model of hypnosis (Jensen et al., 2015). Three categories of factors explained hypnotic responding: biological factors (activity in frontal and anterior cingulate cortices, functional connectivity, structural connectivity, hemisphere asymmetry, and theta band), psychological factors (expectancies, hypnotizability, motivation, and absorption), and social factors (relationship with the therapist and hypnotic context). Several studies have demonstrated the benefits of hypnosis, used alone or in combination with other techniques, on different sides effects of cancer treatments, such as CRF, emotional distress, sleep, pain, and cognitive difficulties (Spiegel, 1985; Cramer et al., 2015; Carlson et al., 2017; Grégoire et al., 2017, 2020a; Arring et al., 2019).

Meditation

Similarly to hypnosis, **meditation** is also considered a non-ordinary state of consciousness. It can be defined as a group of "states, processes, and practices that self-regulate the body and mind, thereby affecting mental and physical events by engaging a specific attentional set" (De Benedittis, 2015). Its aim is to improve voluntary control over mental processes. Different meditation practices exist, with similarities: a quiet location with few distractions, a specific and comfortable posture (sitting or lying down), a focus on attention, and an open attitude of letting thoughts come and go without judgment (Ospina et al., 2008). Phenomenological characteristics of meditation vary according to the type of practice, but alteration of sense of time, space and body representation, along with modifications of emotions and physical sensations are commonly reported (Brandmeyer et al., 2019). The neurophysiological correlates of mediation also depend on the type of meditation and thus are not yet clear (De Benedittis, 2015; Lomas, 2015; Brandmeyer et al., 2019). However, a systematic review showed that mindfulness-based meditation is associated with increased alpha and theta power, while no consistent patterns were observed in beta, delta, and gamma bandwidths. This configuration is indicative of a state of relaxed alertness favoring mental health (Lomas, 2015; Brandmeyer et al., 2019). Meditation has also been shown to positively influence emotional distress, pain, fatigue, and sleep difficulties in patients with cancer (Kim et al., 2013; Carlson et al., 2017; Arring et al., 2019; Ngamkham et al., 2019; Suh et al., 2021). Our study will focus on mindful self-compassion (MSC) meditation (Kotsou and Heeren, 2011; Germer and Neff, 2013, 2019; Neff and Germer, 2013; Neff et al., 2020), which involves a compassionate stance toward oneself when encountering personal difficulties (i.e., self-kindness over self-judgment, sense of common humanity instead of isolation, and mindfulness rather than over-identification), and has been shown to positively influence psychological wellbeing (Kotsou and Leys, 2016; Kılıç et al., 2021). This self-compassionate frame of mind involves being gentle, supportive, accepting, and understanding toward oneself (Germer and Neff, 2019).

Most of these studies about hypnosis and meditation suffer from methodological limitations, the main one being their focus on breast cancer patients only. However, other cancers may have different negative physical and psychological effects. For example, mortality rates vary according to the localization of the tumor (Sung et al., 2021). Additionally, some cancers imply specific adverse events [e.g., prostate cancers (impact on sexuality, possible urinary incontinence) and head and neck cancers (possible tracheotomy with impact on speech)], which may have specific impact on the psychological adaptation to the disease. In addition, most of these studies did not measure the long-term effects of the intervention proposed.

In this study, our aim will be to include a large population of patients with different cancer diagnoses and to rigorously investigate the long-term effects of the interventions. We also want to study another non-ordinary state of consciousness, involving more directly the body: self-induced cognitive trance.

Self-Induced Cognitive Trance

The recent work on hypnosis and meditation has opened the path to study other non-ordinary states of consciousness, such as trance. Self-induced cognitive trance (SICT) is characterized by lucid but narrowed awareness of the environment, hyper-focused immersive experience of flow, enhanced inner imagery, modified somatosensory processing, altered sense of self, and an experience of spiritual travel (Csikszentmihalyi, 1990; Flor-Henry et al., 2017; Grégoire et al., 2021). It is inherited from Mongolian traditional shamanic practice, where is it used in healing interventions (Flor-Henry et al., 2017; Grégoire et al., 2021). However, little is known about the scientifically based phenomenology and neural correlates of SICT. A few case studies report that it is associated with subjective changes in time and space perception, body awareness, thinking and emotional state, with a dissociation and a modulation of perceptions from the environment, as well as decreased pain perception, increased strength, increased sense of happiness, visual imagery, and out-of-body experience (Hove et al., 2016; Flor-Henry et al., 2017; Kawai et al., 2017; Mainieri et al., 2017; Gosseries et al., 2020; Huels et al., 2021). This could be put in parallel with some of the characteristics of hypnosis already described (absorption, dissociation, and modulation of pain, for examples; Vanhaudenhuyse et al., 2014, 2020). Only a few studies used neurophysiological measurements during different kinds of trance, including SICT (Oohashi et al., 2002; Peres et al., 2012; Hove et al., 2016; Flor-Henry et al., 2017; Gosseries et al., 2020), and showed a temporary reconfiguration of brain network architecture. However, results are quite contradictory, and studies are mostly from small samples of trance experts with a lack of well-controlled designs. Moreover, no study has evaluated the clinical applications of SICT, even though it may have potential therapeutic properties, like hypnosis and meditation. Previous works on shamanic trance have reported positive outcomes (e.g., decrease anxiety and increase wellbeing) anecdotally, but not scientifically (Sidky, 2009; Mackinnon, 2012; Grégoire et al., 2021). Thus, investigating SICT's phenomenological and neurophysiological correlates, as well as its clinical applications, and comparing them with hypnosis and meditation interventions seems particularly relevant as it will allow to better understand these non-ordinary states of consciousness and improve available complementary therapeutic options in oncology.

OBJECTIVES

This project's aims are threefold: (1) Evaluating the short- and long-term clinical benefits of hypnosis, meditation, and SICT in terms of CRF, emotional distress, sleep difficulties, and pain (primary outcomes), as well as other psychological variables (e.g., cognitive functioning, adaptation to cancer, psychological flexibility; secondary outcomes) in patients with cancer (through questionnaires); (2) Measuring the evolution of (a) phenomenological and (b) neurobiological correlates of hypnosis, meditation, and SICT in these patients, to better understand the effects of these interventions (through questionnaires and neurobiological measures); (3) Confirming the biopsychosocial model of hypnosis and investigating whether meditation and SICT responsiveness are mediated by the same mechanisms as hypnosis. This study will allow to strengthen the use of hypnosis and meditation for all cancers and to assess the interest of SICT as a new therapeutic complementary option.

METHODS AND ANALYSIS

Design

We designed a longitudinal controlled superiority trial (see **Figure 1**) with 160 patients with cancer who will choose between four conditions (see section "Recruitment" for sample size calculation): hypnosis-based group intervention, MSC meditation-based group intervention, SICT-based group intervention, or a control group. This kind of design has been chosen because patients who receive their preferred therapeutic option seem to be more motivated and exhibit greater adherence to the treatment (King et al., 2005; Preference Collaborative Review Group, 2008; Sedgwick, 2013; Bicego et al., 2021a). Indeed, preference-based trials are increasingly developed as researchers often want to know if an intervention is effective for the participants who choose it, rather than determining the best treatment option irrespective of the participant's choice (Kowalski and Mrdjenovich, 2013). In addition, pragmatic trials are also more and more developed and aim at assessing the effects of interventions in real-life routine practice conditions, rather than under optimal conditions. The results from a pragmatic trial can be generalized and applied in routine practice (Patsopoulos, 2011). Moreover, one of our previous study with oncological patients showed that different intervention groups, based on participants' preferences, were similar at baseline on the sociodemographic, medical, and clinical data investigated (Grégoire et al., 2017). Participants will then know in which group they belong and will be aware of the existence of the other intervention groups.

The structure and intensity of each intervention will be different (see **Figure 1**), because we want to optimize the trainings in a manner that has been shown to be effective in our practice with volunteers and patients (Grégoire et al., 2017, 2018b). More specifically, in the hypnosis-based intervention, patients will participate in 8 weekly sessions (2 h each) during

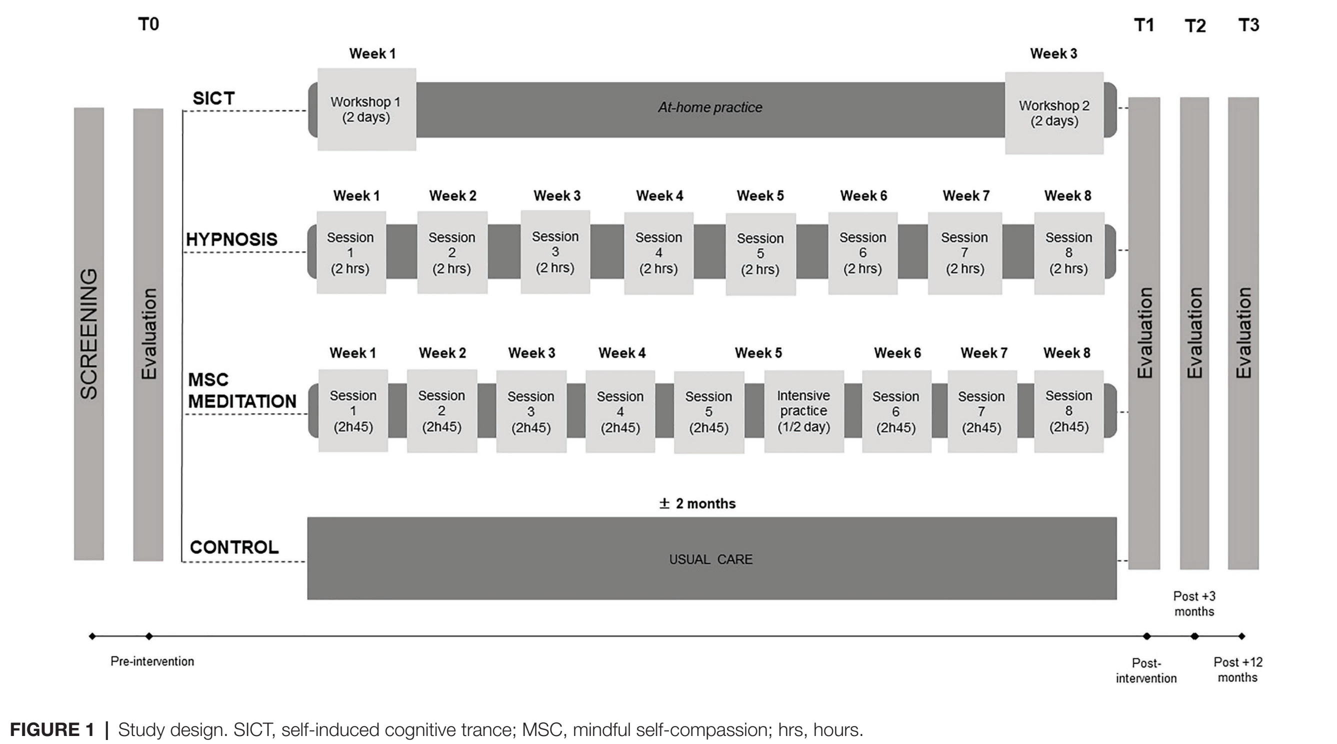


FIGURE 1 | Study design. SICT, self-induced cognitive trance; MSC, mindful self-compassion; hrs, hours.

which they will benefit from guided hypnosis exercises and learn how to implement self-hypnosis in their daily life. In the MSC meditation-based intervention, patients will participate in 8 weekly sessions (2 h45 each) and half a day of intensive practice, in which they will practice different meditation exercises. In the SICT-based intervention, patients will participate in 2-day workshops, spaced by 2 weeks, in which they will learn how to induce SICT, with the use of different sound loops (i.e., electronic designed sounds, voices, and noises). In the three groups, exercises will mostly focus on self-compassion as this component has been shown to be important in oncology settings in previous studies (Grégoire et al., 2017, 2018b, 2020a). At-home practice will also be encouraged between sessions and will be assessed through weekly evaluations (see **Table 1** for measurements used). Participants in the control group will complete the questionnaires and undergo the neurobiological examinations, but they will not attend any of the interventions. They will only benefit from care as usual and have the possibility to participate in one of the three group interventions after the study completion if they want to. More information about the content of each intervention is detailed in section "Interventions".

Each participant from the three intervention groups will be evaluated before the intervention (T0), right after the last group session (T1), at 3 months (T2) and 1 year (T3) post-intervention. Participants in the control group will be evaluated according to a similar schedule (see **Figure 1** for study design, and sections "Procedures" and "Assessments" for more

TABLE 1 | Measurements used in the study.

	Screening	T0	Between T0 and T1 (during the group intervention)	T1	T2	T3
Medical data: diagnosis, time since the end of treatments, psychiatric and neurologic history	X					
VAS: fatigue, distress, pain, sleep difficulties (/10)	X					
No practice of hypnosis, meditation or SICT	X					
Sociodemographic data: age, gender, religious beliefs, professional status, BFI-10		X				
Medical data: time since diagnosis, cancer stage, treatment received		X				
Weekly subjective evaluations (VAS fatigue, emotional distress, sleep difficulties, pain)			X			
Diary of practice (frequency, details about some intense episodes, feelings, impressions, etc.)			X			
Hypnosis/meditation/SICT and imaginary experience: • VAS (/10): expectations, motivation to experience hypnosis/meditation/SICT* • Personal definition and *a priori* according to the chosen intervention (free text)* • CEQ		X				
Quality of life-related variables: • Fatigue and sleep: MFI20 and ISI • Emotional distress: HADS • Pain: VAS (/10) • Self-reported cognitive functioning: FACT-Cog v3 • Other psychological variables: MAC, CERQ, MPFI-24, Empowerment [7 VAS (/10)], major life events		X		X	X	X
Neurophysiological and biological variables:						
• EEG resting state		X		X		X
• EEG during hypnosis, meditation or SICT state*				X		X
• ECG, EMG, and respiration (effort and pressure)		X		X		X
• Core body temperature		X		X		X
• Tumor marker rates		X		X		X
Questionnaires about the intervention: • Free recall of a hypnosis/meditation/SICT/intense autobiographical episode (written, according to the intervention) • MCQ • MEQ30 • Presence questionnaire • VAS (/10): Quality of the relationship with the therapist*				X	X	X

*BFI, Big Five Inventory; CEQ, Creative Experiences Questionnaire; CERQ, Cognitive Emotion Regulation Questionnaire; FACT-Cog, Functional Assessment of Cancer Therapy – Cognitive Function; HADS, Hospital Anxiety and Depression Scale; ISI, Insomnia Severity Index; MAC, Mental Adjustment to Cancer Scale; MCQ-30, Memory Characteristics Questionnaire; MEQ30, Revised Mystical Experience Questionnaire; MFI-20, Multidimensional Fatigue Inventory; MPFI-24, Multidimensional Psychological Flexibility Index; ECG, electrocardiogram; EEG, electroencephalogram; EMG, electromyogram; VAS, Visual Analogue Scale. *Only for the participants in one of the 3 intervention groups.*

information on procedures and measurements). The first intervention group (hypnosis) started in April 2021.

Eligibility

Inclusion criteria will be: to be at least 18 years old, no psychiatric or neurological disorders, no current and regular practice of hypnosis, meditation or trance, diagnosis of cancer (any localization except brain tumors to avoid any effect on EEG data or severe cognitive impairments that could prevent filling the questionnaires adequately), and completion of active treatments (surgery, chemotherapy, radiotherapy) for less than a year. Interested patients will be screened for fatigue, emotional distress, sleep difficulties, and pain (VAS score ≥ 4/10 for one of these four symptoms for inclusion), based on recommendations from our past studies (Grégoire et al., 2018a,b).

Recruitment

Participants will be recruited over a 2- to 3-year period, mainly at the University Hospital of Liège (Belgium) but also in other structures and through social medias. Potentially eligible participants will be identified in various ways. First, different health care professionals working with oncological patients (e.g., oncologists, radiotherapists, psychologists, nurses) have been informed of the study and asked to talk about it to their patients who have completed their treatment. In the hospital, interested patients will be advised to contact us by phone or e-mail. Second, flyers and posters will be displayed in several strategic areas in and outside the hospital (e.g., oncology, radiotherapy, algology services' waiting rooms, doctors' office, private consultations cabinet, other hospitals, health associations, health professionals in Belgium or France), which allow other health professionals and patients to be informed about our study. Third, we will use social medias (e.g., personal and professional Facebook accounts, Mewe) to inform as much people as possible about our study.

The recruitment started in December 2020 and is ongoing. Sample size has been determined by a power analysis allowing a comparison between the four groups. The sample size calculation was based on a repeated measures ANOVA ("between factors"). Alpha was set at 0.05, power at 90%, the standardized effect size at 0.3 and the expected correlation between repeated measures at 0.6. These coefficients were successfully used in our previous studies (Grégoire et al., 2017, 2020a). As there will be four groups and four measurements times, a total of 116 patients will be needed according to this analysis. Based on our previous experience with a similar design, we expect a 20% drop-out rate, leading to a sample of 140 patients (35 in each condition; Grégoire et al., 2020a). As our group sessions will include approximately 10 persons, we will then aim at recruiting 40 participants in each condition (i.e., four groups of 10 participants in hypnosis, MSC meditation, and SICT groups and 40 participants in the control group), for a final sample of 160 participants. Participants will have the opportunity to choose the intervention in which they want to participate. Most of the participants already have a strong preference for one of the three interventions when contacting us. However, each interested person who fill the inclusion criteria will be asked to watch a short video we made to explain the three interventions and the aims of the study.[1] This allows every participant to receive standardized information about the study and to make or confirm their choice of intervention. However, once 40 participants have been recruited in a condition, it will not be possible for other interested persons to be included in this condition. These persons will be proposed to integrate the control group or another active condition.

Procedures

A written consent will be obtained from each participant at the beginning of the study. The investigators will remain available by phone or e-mail during the study duration to answer any question.

Screening

During the first telephone contact with the interested participants, they will be informed of the protocol and study design. If interested, one of the two main investigators (CG or NM) will ensure that participants meet all the inclusion criteria. If eligible, they will have to watch the information video and confirm their choice of intervention. Once a group is almost complete (i.e., around 10 participants), the first evaluation (T0) will be scheduled.

Measurement Points (T0, T1, T2, and T3)

Each of these four measurement points will be completed by every participant and will consist of several questionnaires (at T0, T1, T2, and T3) and neurobiological measures (at T0, T1, and T3; see **Table 1**). Questionnaires will be completed online at home, on a secured website.[2] For those who do not have a computer or are not comfortable using one, it will be possible to make an appointment with one of the investigators to complete the questionnaires with them. At each measurement point, the risk of missing responses is very low, as the online questionnaires are designed in a way that does not allow for missing answers. The investigators will check before the first group session that all participants have answered the questionnaires and remind them if it is not done. For the participants who will complete the questionnaires with one of the investigators, all answers will be checked to minimize missing data.

The neurophysiological and biological measures will be collected at the hospital (CHU of Liège), during a session of ±1 h (see point 3.6.). Tumor marker rates will be collected directly from the patient's medical record or by asking them to send us their last blood test result. All data will be anonymized: a code will be attributed to each participant and used during the entire study. Only the researchers involved in this study will have access to the final datasets. Review of the trial process and difficulties encountered are discussed regularly between the researchers involved.

[1] https://www.youtube.com/watch?v=lRhs8jK-RCE
[2] www.alchemer.com

Interventions

Hypnosis

This intervention will include eight 2h weekly sessions in groups of approximately 10 participants. This protocol has been developed and will be led by one of the authors (MEF), who is an anesthesiologist and an international expert in hypnosis. A large part of the first session will be devoted to answering to the participants' questions and giving information about hypnosis. First exercises will be proposed as introduction to hypnosis: a focused listening of different musical records and some imagery exercises. The aim is for the participants to notice their abilities to access mental imagery and hypnosis. During the other sessions, a debriefing of at-home practice will be proposed, along with different hypnosis exercises. They will be repeated and discussed, and participants will also attempt to induce hypnosis by themselves during the last session. **Table 2** summarizes the exercises that will be done at each session. Between sessions, participants will be encouraged to practice the different exercises at home with the help of audio recordings of the hypnosis exercises. This is essential to take full advantage of hypnosis without the help of a therapist. A self-care approach will be fostered during the intervention, more specifically during the debriefings of the participants' practice, if they encounter difficulties or ruminations. Concrete strategies to deal with such difficulties will be proposed by the therapist. The importance to take time for themselves will be emphasized, as well as to allow themselves to take a break from their routine and practice hypnosis in their daily life to improve their wellbeing, in an adaptation of our usual practice with oncological patients in clinical and experimental settings (Charland-Verville et al., 2017; Grégoire et al., 2018b, 2020a).

MSC Meditation

This intervention will include eight sessions of 2 h45 in groups of approximately 10 participants. It will be led by two of the authors (VvN and SdR), recognized as experts in MSC meditation. The intervention is based on the practice of MSC meditation, already described above. At the beginning of the first session, a presentation of the intervention will be proposed along with explanations about meditation and answers to participants' questions. Each session will focus on a theme, linked to self-compassion (see **Table 2**), and will be based on formal practices (e.g., guided meditation), experiential exercises, informal practices linked to the daily life and theoretical information about compassion and emotions. Each exercise will be debriefed in small or large groups, in order to allow each participant to share their feelings, thoughts and questions. Between each session, individual practice will be encouraged and audio recordings of different exercises will be given to the participants.[3] Between the fifth and the sixth session, half a day of silent intensive practice will be organized, during which all the exercises previously worked on will be reviewed and deepened.

[3]https://soundcloud.com/emergences/sets/cycle-auto-compassion

Self-Induced Cognitive Trance

This intervention will include two 2-day workshops, spaced by two weeks, in groups of approximately 10 participants. A part of the first session will be devoted to explain what is SICT, how it can be accessed and how the workshops will take place. Possible questions among the participants will be discussed. To induce SICT, different sound loops (i.e., electronic designed sounds, voices, and noises, inspired by the ones produced during traditional shamanic rituals) will be used. They were designed and previously employed by one of the authors (CS), who is the first Westerner to become an *ugdan* (female shaman in Mongolia), and recognized as an international expert on SICT (Sombrun, 2006, 2012). Different kinds of trances will be taught during the workshops: trances with intention, during which participants follow an intention [e.g., refocusing on oneself ("anchor trance"), finding one's place and being assertive ("territorial trance"), reconnecting with one's inner power ("power trance")], and trances without intention ("free trances"), during which participants let themselves and their body get carried away by the experience of trance. During each session, 4–5 SICT exercises will be proposed (see **Table 2**). Typically, participants will lie down and listen to one of the sound loops (± 30 min). The access to SICT can manifest itself through various elements: vocalizations or movements for examples. Every exercise will be debriefed in group, and a focus on self-compassion will be suggested during the workshops using specific intentions (e.g., an intention to be gentle with oneself, set self-boundaries). At the end of the two workshops, all participants will have found a way to induce SICT without the help of the sound loops. At-home practice will be encouraged during the two weeks between the workshops and after them.

The three interventions will take place in Liège (hypnosis, SICT) and Brussels (MSC meditation). During the entire study duration, each participant will benefit from usual care, including medical care, oncological revalidation, and individual psychological help if necessary. Although no adverse event has been reported in our previous studies on hypnosis with patients with cancer or chronic pain (Grégoire et al., 2017, 2018b, 2020a; Bicego et al., 2021b) and clinical practice, it could be possible that a patient feels uncomfortable during the exercises. In this case, they will have the possibility to stop the session and their participation in the study. The therapist or the experimenter can also propose a meeting to discuss their difficulties and, if necessary, suggest a meeting with a psychologist or physician with whom we collaborate. Any reason for drop-out will be consigned. Adherence to the study is fostered by the facts that all participants are volunteers and motivated to participate and that the assessments can be mainly done at home and require only three travels to the hospital during the whole study. Several reminders will also be sent to the participants who forget to answer the questionnaires in time and who did not ask to quit the study.

Assessments

Table 1 summarizes the different parameters used at each measurement point.

TABLE 2 | Summary of the content of the sessions for each intervention.

	Hypnosis	MSC meditation		Self-induced cognitive trance (SICT)
Session 1	Explanations about hypnosis and answers to questions. Exercises based on focused listening of musical records specially composed by a certified music therapist, and mental imagery.	Explanations about meditation and answers to questions. Exercises based on the discovery of mindful self-compassion (soothing touch, self-compassion break).	Workshop 1—Day 1	Explanations about SICT and answers to questions. SICT exercises using sound loops. Individual search of one or several movements or sounds to induce SICT without listening to the sound loops. Debriefing of each exercise in group. Focus on the transformative power of SICT (i.e., impact on bodily sensations and somatosensory processes, on cognition and emotions).
Session 2	Discussion of self-care strategies in accordance to the difficulties encountered by the participants. Soothing White Clouds exercises.	Mindful self-compassion exercises (affectionate breathing, grounding, mindfulness in daily life, present moment).		
Session 3	Discussion of self-care strategies in accordance to the difficulties encountered by the participants. Safe Place exercises.	Benevolent love exercises (affectionate breathing, love for a close one, self-compassion, and benevolent love).	Workshop 1—Day 2	SICT exercises using different sound loops and individual movements or sounds. Debriefing of each exercise in group. Focus on the transformative power of SICT.
Session 4	Discussion of self-care strategies in accordance to the difficulties encountered by the participants. Garden of dreams exercises.	Discovery of one's compassionate voice exercises (self-compassion, security and self-criticism, self-compassionate letter).		
Session 5	Discussion of self-care strategies in accordance to the difficulties encountered by the participants. Pain and Colors exercises.	Importance of living intensely (giving and receiving compassion, personal values and wishes, silver lining, compassionate listening).	Workshop 2—Day 1	Debriefing of the 2-week individual practice. SICT exercises using different sound loops and individual movements or sounds. Debriefing of each exercise in group. Consolidation of the participants' self-induction ability. Focus on the interactions during SICT (i.e., with the environment).
Between sessions 5 and 6	/	Half a day of silent intensive practice (review and deepening of all the exercises previously worked on)		
Session 6	Discussion of self-care strategies in accordance to the difficulties encountered by the participants. Glove Analgesia exercises.	Dealing with difficult emotions [self-compassion, acceptation and management of emotions, "soften, soothe, allow" (basic and specific to shame)].		
Session 7	Discussion of self-care strategies in accordance to the difficulties encountered by the participants. Levitation exercises.	Exploration of difficult relationships (compassionate friend, unfulfilled needs, Qi gong, self-compassion break in relationships, compassion with equanimity, dealing with compassion fatigue).	Workshop 2—Day 2	Debriefing of the 2-week individual practice. SICT exercises using different sound loops and individual movements or sounds. Debriefing of each exercise in group. Consolidation of the participants' self-induction ability. Focus on the interactions during SICT.
Session 8	Discussion of self-care strategies in accordance to the difficulties encountered by the participants. Stories and Metaphors exercises.	Embracing life (self-compassion and compassion for others, appreciation and gratitude, self-compassion bracelets).		
Each session	Debriefing of the previous session and at-home practice.			

Questionnaires

Approximate duration of questionnaires completion is 40–50 min at each measurement point.

General Information

Sociodemographic and medical data (gender, age, education level, professional activity, spiritual believes, personal history of cancer and treatment, possible recurrence during the study) will be collected. The Big Five Inventory (short version—10 items; Courtois et al., 2020) will also be administered. This questionnaire aims at rapidly assessing personality according to the 5 domains of the original Big Five Inventory (extraversion, conscientiousness, negative emotionality, open-mindedness, and agreeableness).

Physical and Psychological Functioning

- *Visual Analogue Scales (VAS)*: Different VAS will assess the participants' emotional state (pain, empowerment, fatigue, emotional distress, sleep difficulties). VAS will also be used during the duration of the group intervention to assess the severity of fatigue, distress, sleep difficulties, and pain on a weekly basis. Each score will be comprised between 0 and 10.
- *Multidimensional Fatigue Inventory* (MFI-20; Smets et al., 1995): This 20-item scale is designed to measure fatigue on 5 different dimensions: general fatigue, physical fatigue, mental fatigue, reduced motivation, and reduced activity.

- *Insomnia Severity Index* (ISI; Savard et al., 2005): This 7-item questionnaire investigates the sleep complaints and the associated distress.
- *Hospital Anxiety and Depression Scale* (HADS; Zigmond and Snaith, 1983): This 14-item scale measures anxiety (7 items) and depression (7 items) and has been validated for people with somatic illnesses.
- *Mental Adjustment to Cancer Scale* (MAC; Watson et al., 1988): This questionnaire assesses the coping styles and adjustment to cancer. It is divided into two sub-scales: positive and negative adjustments.
- *Cognitive Emotion Regulation Questionnaire* (CERQ; Garnefski et al., 2001): This multidimensional scale investigates the cognitive emotion regulation strategies used after experiencing negative events linked to the disease or its treatments.
- *Multidimensional Psychological Flexibility Index* (MPFI-24; Grégoire et al., 2020b): This 24-item scale assesses the psychological flexibility and psychological inflexibility, through 12 dimensions (6 linked to psychological flexibility: acceptance, present moment awareness, self as context, defusion, contact with values, committed actions; and 6 for psychological inflexibility: experiential avoidance, lack of present moment awareness, self as content, fusion, lack of contact with values, inaction) and 2 total scores (psychological flexibility and inflexibility).
- *Life events*: To assess whether the participants experienced any major life event during or after the intervention, we will ask them the following question (yes/no): "*Did you experience any major life event since the last time you completed these questionnaires* (wedding, birth, job loss, accident, health problem, etc.)."

Cognitive Functioning

- *Functional Assessment of Cancer Therapy – Cognitive Function* (FACT-Cog v.3; Joly et al., 2012): This 20-item questionnaire measures the participant's subjective cognitive functioning over the past week.

Questionnaires Linked to the Chosen Intervention (Hypnosis, Meditation or SICT)

- *VAS (/10)* concerning the expectations and motivation to participate in the intervention (T0), then concerning the quality of the relationship with the therapist (T1–T3).
- *Questions about the chosen intervention (free text)*: This will aim at investigating the personal definition of hypnosis, meditation, or SICT of participants, their beliefs and *a priori* concerning the intervention they will participate in.
- *Creative Experiences Questionnaire* (CEQ; Merckelbach et al., 2001): This 25-item (true/false) scale has been designed to investigate the fantasy proneness. Questions are about, for examples, imaginative friend or animal, belief in supernatural beings, tendency to dream of fantasize, or out-of-body experiences.
- *Free recall of an intense hypnosis/meditation/SICT episode (intense autobiographical episode for the participants in the control group)*: This will be used to collect phenomenological data about each non-ordinary state of consciousness involved in the study. As patients in the control group will not participate in any of the interventions, they will have to describe an important autobiographical memory that happened during the study (e.g., wedding, birth). The recall of the episode will have to be as detailed as possible.
- *Memory Characteristics Questionnaire* (MCQ; Johnson et al., 1988): This 16-item questionnaire is adapted from the original questionnaire. It assesses a wide range of memory characteristics (visual details, complexity, spatial, and temporal information, feelings). Participants will be asked to answer the questionnaire while thinking about one specific memory (the one used in the free recall).
- *Revised Mystical Experience Questionnaire* (MEQ30; Maclean et al., 2012): This 30-item scale is designed to assess the mystical characteristics of a specific episode or memory (in this study, a hypnosis, meditation, or SICT episode), across four factors: mystical, positive mood, transcendence of time/space, and ineffability.
- *Presence questionnaire* (Heck et al., 2021): This 12-item scale has been initially designed to measure the feeling of presence after the use of virtual reality. As the original questionnaire was focused on the virtual environment specific to virtual reality, we adapted it to a more general "suggested" environment, such as in hypnosis, meditation, and SICT. Questions are about the impression to have been present in the environment, to have interacted with it, or to have felt the presence of other people in it, for examples.

Neurophysiological and Biological Measures

We will record an electroencephalogram (EEG) resting state (256 EGI, Geodesics) during 15 min, combined (T1 and T3) or not (T0) with an EEG during hypnosis, meditation, or SICT state (15 min). During each EEG, we will also record an electrocardiogram (ECG) and an electromyogram (EMG), and measure the breathing (effort and pressure, Polygraph Input Box of EGI system) and body temperature. Tumor marker rates (according to the patient's diagnosis) will also be collected in the patient's medical record or based on their blood test results.

Data Coding and Storage

Most questionnaires will be completed electronically (on www.alchemer.com) by the participants, through a secured link sent by the experimenters. Data encoding will be done automatically by Alchemer and will be regularly checked by the experimenters. Final databases and manually coded data (e.g., from manually completed questionnaires or directly collected from medical records) will be stored on a protected server from the University, protected by a password. Data coding and storage comply with the General Data Protection Regulation (GDPR).

Statistical Analyses

Baseline (T0) demographic, medical, and psychological data will be compared between groups to test for initial groups equivalency using inferential statistics, including ANOVAs and chi-square tests. Normality of data will be checked using Shapiro–Wilk test. Group-by time changes, and pre- and post-assessment comparison of each variable within each group will be assessed using repeated measures MANOVA followed by post hoc comparisons, on the participants who completed all the needed assessments points. Hierarchical linear regressions will be conducted to investigate the factors associated with the evolution of our main outcomes. All tests will be two-tailed, and the alpha will be set at 0.05. For EEG data analysis, MATLAB, EEGLAB, FieldTrip, Brainstorm, and in-house MATLAB and Python modules will be used. EEG signal will be preprocessed as in our previous works [downsampling, epoching, filtering, removing noise (notably using ICA), channels interpolation, re-referencing; Carrière et al., 2020; Thibaut et al., 2021]. EEG markers will then be extracted: spectral (e.g., power spectrum density), connectivity (e.g., weighted symmetrical mutual information), complexity (e.g., permutation entropy; Engemann et al., 2018). We will also apply graph theory connectivity measures to further assess changes between brain regions (Chennu et al., 2017; Carrière et al., 2020; Thibaut et al., 2021). Results will be corrected for multiple comparisons and considered significant at $p < 0.05$. Statistical analyses will be performed after all data have been collected for T2, then for T3 and T4.

All participants will be informed by e-mail about the final results of the study. Scientific publications and presentations will also be planned.

DISCUSSION

Based on previous studies showing the benefits of different mind–body interventions, such as hypnosis and meditation, on cancer patients' quality of life (Spiegel, 1985; Kim et al., 2013; Cramer et al., 2015; Carlson et al., 2017; Grégoire et al., 2017, 2020a; Ngamkham et al., 2019; Suh et al., 2021), and on the growing interest of these patients for such interventions, we designed a longitudinal controlled study assessing the benefits of three mind–body group interventions based on non-ordinary states of consciousness: hypnosis, MSC meditation, and SICT. Our aim is to investigate the short- and long-term effects of these interventions on physical and psychological symptoms of post-treatment oncological patients and to investigate their phenomenological experiences, their neurophysiological correlates, and their mechanisms of action. Our wish is also to include a variety of cancer diagnoses to address a major gap in the scientific literature: the lack of data about the effects of psychosocial interventions on cancers other than breast cancers.

Based on the existing scientific literature, our previous studies, and preliminary data (not published), we made several hypotheses:

1. **Hypnosis, MSC meditation, and SICT will have positive overall benefits on the quality of life of cancer patients**, with increased effects over time. More specifically, we expect that the three interventions will have superior positive effects on the patients' quality of life than the control group. We expect that these effects will be significant right after the intervention (T1) and will stay stable or increase at 3- and 12-month follow-up (T2 and T3) due to continued practice. Moderate difference might be observed between the three intervention groups. For example, SICT could impact more the body-related variables (e.g., pain and fatigue) as it involves the body more directly (e.g., in the induction process). We expect that a small improvement will be achieved in the control group as well, but much weaker than in the three intervention groups.
2. **Phenomenological experiences will change over time and will be intervention-dependent**. We hypothesize that over time and in the three intervention groups, phenomenology will become more joyful and richer. As the phenomenology of hypnosis, MSC meditation, and SICT has been very little studied, we do not have any specific hypothesis concerning the phenomenological differences between the three interventions. Our analyses are exploratory. We however expect that the phenomenology will change in each intervention group between the different measurement times, as the participants are expected to continue to practice regularly after the intervention. We do not expect any major evolution of the phenomenological experiences in the control group.
3. **Neurophysiological changes will be different between the three intervention groups**. We hypothesize that neurophysiological changes will vary according to the intervention and that they will evolve over time, in response to the continued practice of the participants after the intervention. Previous studies showed increased theta power and changes in gamma activity during hypnosis (Vanhaudenhuyse et al., 2020) and increased alpha and theta power, during meditation (Lomas, 2015; Brandmeyer et al., 2019). We hypothesize to observe the same patterns in our study. We do not expect any evolution of neurophysiology in the control group. We also expect differences in neurophysiological correlates of resting state, depending on patients' clinical outcomes, especially long-term ones. For example, we could expect slower activity if patients are feeling less anxious.
4. **We will be able to confirm, at least in part, the biopsychosocial model of hypnosis**, and MSC meditation and SICT responsiveness mechanisms will be in part similar to the ones implied in hypnosis responsiveness. We also expect other potential factors (linked to body physiology: temperature, ECG, EMG, breathing) to correlate with hypnosis, MSC meditation, and/or SICT responsiveness.

Participants are not aware of the specific hypotheses of our study. However, they have a general idea of the study aim, as the flyers and posters used for the recruitment underlined the fact that the purpose of each intervention is to improve their quality of life.

There are some limitations of our study. First, only patients who have finished their active treatments for less than a year can participate in the study. It is possible that patients who finished their treatments longer ago, or who are still in treatment, will be willing to participate. The same limitation applies for most of our inclusion criteria: some patients with brain tumor, or who are currently practicing hypnosis or meditation for example, could be interested in the study, and will not be able to participate. It is then possible that these inclusion criteria will impact the recruitment process and our results. However, they were chosen in order to minimize the baseline differences in our sample. Another limitation is linked to the profile of our participants. As our study takes place in Belgium, our participants will mainly come from Belgium or France or maybe French-speaking areas of other bordering countries. This means that all our participants will come from western and industrialized areas, which could induce some similarities among them. Our three interventions are also proposed in a group setting, which could discourage some people to participate as they could be uncomfortable in sharing their experience and thoughts with others. In addition, the therapists involved in our study could also be considered as a potential bias. Indeed, some of them are recognized as international experts in their discipline, and the fact that they will lead the group sessions is likely to impact the recruitment process, with more patients wanting to participate in their interventions. Finally, the design of our study could be considered as a limitation. First, our study will not be randomized, even though randomized-controlled trials are generally considered the gold standard for research aiming at assessing the effectiveness of an intervention (Hariton and Locascio, 2018). This preference-based design was chosen because we did not find it relevant for a patient who wants to participate, for example, in the hypnosis intervention to be obliged to participate in another intervention. Indeed, patients who can chose their treatment have more motivation and a greater adherence (King et al., 2005; Preference Collaborative Review Group, 2008; Sedgwick, 2013; Bicego et al., 2021a). We also chose this design based on our clinical observations. After a generally long cancer journey, where patients rarely had the opportunity to make any choice and had to endure several intense treatments and their side effects, it seemed important to us to allow them to finally take back the control over their care and choose the intervention they want to participate in. In addition, preference-based and pragmatic trials are more and more represented in scientific studies when the aim is to assess the effectiveness of an intervention for the participants who chose it, rather than determining the best treatment option (Kowalski and Mrdjenovich, 2013), or to assess the effectiveness of the intervention in real-life conditions (Patsopoulos, 2011). Second, the designs of the three interventions are different in terms of duration and frequency. As explained above, we decided to use these designs because they have been shown to be acceptable and effective in our practice with volunteers and patients. As the main aim of this study is not to investigate how each intervention allows an improvement of the patients' quality of life, using different intervention designs seems understandable and acceptable. If our results are conclusive, future research should continue to investigate these three interventions and their mechanisms of action by standardizing their designs.

In conclusion, this project is particularly original, timely relevant and innovative. First, it will allow to adequately evaluate the short- and long-term effects of the interventions to improve the quality of life of a mixed population of patients who often suffer from intense and debilitating long-lasting cancer-related symptoms, along with behavioral, phenomenal and neurophysiological changes. This could lead to an improvement of mind–body interventions in oncology settings in the future, along with a better understanding of hypnosis, meditation, and SICT phenomena and their impact on quality of life. We could also imagine proposing these mind–body interventions to other clinical populations in the future (e.g., patients with chronic pain). Second, combining phenomenology and electrophysiology with rigorous methodology will be an asset to dive deep into the cognitive and brain functions related to hypnosis, meditation, and SICT, including their impact on physical and psychological symptoms.

ETHICS STATEMENT

The studies involving human participants were reviewed and approved by Hospital-Faculty Ethics Committee of Liège. The patients/participants provided their written informed consent to participate in this study.

AUTHOR CONTRIBUTIONS

CG, AV, and OG participated in the conception and design of the study and in drafting the manuscript. CS, MEF, IK, and GJ participated in the conception and design of the study and in revising the manuscript critically for important intellectual content. NM, VvN, SdR, and SL participated in revising the manuscript critically for important intellectual content. CS, MEF, VvN, and SdR also led the intervention groups. All authors have read and approved the final manuscript and agreed to be accountable for all aspects of the work in ensuring that questions related to the accuracy or integrity of any part of the work are appropriately investigated and resolved.

FUNDING

This study was funded by the Belgian National Funds for Scientific Research (FRS-FNRS)—Télévie (Grant 7460320), the Belgian Cancer Foundation (Grant C/2020/1357), the University of Liège and University Hospital of Liège, and the Benoit Foundation (Brussels—Belgium). These funds financed the different researchers involved in this study. The funding bodies have no role in the design of the study and collection, analysis, and interpretation of data and in writing the manuscript. CG is a postdoctoral researcher, OG is a Research Associate, and SL is a Research Director at the F.R.S-FNRS.

ACKNOWLEDGMENTS

We are grateful to all the patients who participated and will participate in the study. We also thank all the health professionals working at the CHU of Liège or elsewhere in Belgium and France who are helping us to recruit the participants needed for our study, and Alain Collinet, who composed the musical records used in the first session of the hypnosis groups. Finally, we want to thank the BIAL Foundation and the Mind Care Foundation for their support.

REFERENCES

Achimas-Cadariu, P., Iancu, M., Pop, F., Vlad, C., and Irimie, A. (2015). Psychological screening and health related quality of life in Romanian breast cancer survivors. *J. Evid.-Based Psychother.* 15, 267–284.

Arring, N. M., Barton, D. L., Brooks, T., and Zick, S. M. (2019). Integrative therapies for cancer-related fatigue. *Cancer J.* 25, 349–356. doi: 10.1097/PPO.0000000000000396

Batty, G. D., Russ, T. C., Stamatakis, E., and Kivimäki, M. (2017). Psychological distress in relation to site specific cancer mortality: pooling of unpublished data from 16 prospective cohort studies. *BMJ* 356:j108. doi: 10.1136/bmj.j108

Bicego, A., Monseur, J., Rousseaux, F., Faymonville, M.-E., Malaise, N., Salamun, I., et al. (2021a). Drop-out from chronic pain treatment programmes: is randomization justified in biopsychosocial approaches? *J. Rehabil. Med.* 53:jrm00185. doi: 10.2340/16501977-2824

Bicego, A., Rémy, H., Diep, A. N., Donneau, A.-F., Faymonville, M.-E., Nyssen, A.-S., et al. (2021b). Psychological interventions influence patients attitudes and beliefs about their chronic pain a 6 month follow up. Chronic Pain & Management. Available at: https://www.gavinpublishers.com/articles/research-article/Chronic-Pain-Management/psychological-interventions-influence-patients-attitudes-and-beliefs-about-their-chronic-pain-a-6-month-follow-up (Accessed April 29, 2021).

Bicego, A. Y., Rousseaux, F., Faymonville, M.-E., Nyssen, A.-S., and Vanhaudenhuyse, A. (2021c). Neurophysiology of hypnosis in chronic pain: a review of recent literature. *Am. J. Clin. Hypn.* 64, 62–80. doi: 10.1080/00029157.2020.1869517

Brandmeyer, T., Delorme, A., and Wahbeh, H. (2019). The neuroscience of meditation: classification, phenomenology, correlates, and mechanisms. *Prog. Brain Res.* 244, 1–29. doi: 10.1016/bs.pbr.2018.10.020

Carlson, L. E., Zelinski, E., Toivonen, K., Flynn, M., Qureshi, M., Piedalue, K.-A., et al. (2017). Mind-body therapies in cancer: what is the latest evidence? *Curr. Oncol. Rep.* 19, 67. doi: 10.1007/s11912-017-0626-1

Carrière, M., Cassol, H., Aubinet, C., Panda, R., Thibaut, A., Larroque, S. K., et al. (2020). Auditory localization should be considered as a sign of minimally conscious state based on multimodal findings. *Brain Commun.* 2:fcaa195. doi: 10.1093/braincomms/fcaa195

Charland-Verville, V., Faymonville, M.-E., Vanhaudenhuyse, A., Raaf, M., Grégoire, C., and Bragard, I. (2017). Apprentissage de l'autohypnose/autobienveillance en oncologie. Pour qui? Comment? Dans quel intérêt? Une revue de la littérature internationale. *Psycho-Oncologie* 11, 51–55. doi: 10.1007/s11839-017-0614-6

Chennu, S., Annen, J., Wannez, S., Thibaut, A., Chatelle, C., Cassol, H., et al. (2017). Brain networks predict metabolism, diagnosis and prognosis at the bedside in disorders of consciousness. *Brain* 140, 2120–2132. doi: 10.1093/brain/awx163

Courtois, R., Petot, J.-M., Plaisant, O., Allibe, B., Lignier, B., Réveillère, C., et al. (2020). Validation française du Big Five Inventory à 10 items (BFI-10). *L'Encéphale* 46, 455–462. doi: 10.1016/j.encep.2020.02.006

Cramer, H., Lauche, R., Paul, A., Langhorst, J., Kümmel, S., and Dobos, G. J. (2015). Hypnosis in breast cancer care: a systematic review of randomized controlled trials. *Integr. Cancer Ther.* 14, 5–15. doi: 10.1177/1534735414550035

Csikszentmihalyi, M. (1990). *Flow: The Psychology of Optimal Experience*. New-York, NY, USA: Harper & Row.

De Benedittis, G. (2015). Neural mechanisms of hypnosis and meditation. *J. Physiol. Paris* 109, 152–164. doi: 10.1016/j.jphysparis.2015.11.001

De Benedittis, G. D. (2021). Neural mechanisms of hypnosis and meditation-induced analgesia: a narrative review. *Int. J. Clin. Exp. Hypn.* 69, 363–382. doi: 10.1080/00207144.2021.1917294

de Vries, M., and Stiefel, F. (2014). "Psycho-oncological interventions and psychotherapy in the oncology setting," in *Psycho-Oncology*. ed. U. Goerling (Berlin, Heidelberg: Springer Berlin Heidelberg), 121–135.

Demertzi, A., Soddu, A., Faymonville, M.-E., Bahri, M. A., Gosseries, O., Vanhaudenhuyse, A., et al. (2011). Hypnotic modulation of resting state fMRI default mode and extrinsic network connectivity. *Prog. Brain Res.* 193, 309–322. doi: 10.1016/B978-0-444-53839-0.00020-X

Demertzi, A., Vanhaudenhuyse, A., Noirhomme, Q., Faymonville, M.-E., and Laureys, S. (2015). Hypnosis modulates behavioural measures and subjective ratings about external and internal awareness. *J. Physiol. Paris* 109, 173–179. doi: 10.1016/j.jphysparis.2015.11.002

Die Trill, M. (2013). Anxiety and sleep disorders in cancer patients. *EJC Suppl.* 11, 216–224. doi: 10.1016/j.ejcsup.2013.07.009

Dodd, M. J., Cho, M. H., Cooper, B. A., and Miaskowski, C. (2010). The effect of symptom clusters on functional status and quality of life in women with breast cancer. *Eur. J. Oncol. Nurs.* 14, 101–110. doi: 10.1016/j.ejon.2009.09.005

Eliott, J. A., Kealey, C. P., and Olver, I. N. (2008). (Using) complementary and alternative medicine: the perceptions of palliative patients with cancer. *J. Palliat. Med.* 11, 58–67. doi: 10.1089/jpm.2007.0102

Elkins, G. R., Barabasz, A. F., Council, J. R., and Spiegel, D. (2015). Advancing research and practice: the revised APA division 30 definition of hypnosis. *Int. J. Clin. Exp. Hypn.* 63, 1–9. doi: 10.1080/00207144.2014.961870

Engemann, D. A., Raimondo, F., King, J.-R., Rohaut, B., Louppe, G., Faugeras, F., et al. (2018). Robust EEG-based cross-site and cross-protocol classification of states of consciousness. *Brain* 141, 3179–3192. doi: 10.1093/brain/awy251

Flor-Henry, P., Shapiro, Y., and Sombrun, C. (2017). Brain changes during a shamanic trance: altered modes of consciousness, hemispheric laterality, and systemic psychobiology. *Cogent Psychol.* 4:1313522. doi: 10.1080/23311908.2017.1313522

Garnefski, N., Kraaij, V., and Spinhoven, P. (2001). Negative life events, cognitive emotion regulation and emotional problems. *Personal. Individ. Differ.* 30, 1311–1327. doi: 10.1016/S0191-8869(00)00113-6

Germer, C. K., and Neff, K. D. (2013). Self-compassion in clinical practice: self-compassion. *J. Clin. Psychol.* 69, 856–867. doi: 10.1002/jclp.22021

Germer, C., and Neff, K. (2019). "Mindful self-compassion (MSC)," in *Handbook of Mindfulness-Based Programmes*. ed. I. Ivtzan (Routledge), 357–367.

Gonzalez, B. D., Eisel, S. L., Qin, B., Llanos, A. A. M., Savard, J., Hoogland, A. I., et al. (2021). Prevalence, risk factors, and trajectories of sleep disturbance in a cohort of African-American breast cancer survivors. *Support Care Cancer* 29, 2761–2770. doi: 10.1007/s00520-020-05786-2

Gosseries, O., Fecchio, M., Wolff, A., Sanz, L. R. D., Sombrun, C., Vanhaudenhuyse, A., et al. (2020). Behavioural and brain responses in cognitive trance: a TMS-EEG case study. *Clin. Neurophysiol.* 131, 586–588. doi: 10.1016/j.clinph.2019.11.011

Götze, H., Friedrich, M., Taubenheim, S., Dietz, A., Lordick, F., and Mehnert, A. (2020). Depression and anxiety in long-term survivors 5 and 10 years after cancer diagnosis. *Support Care Cancer* 28, 211–220. doi: 10.1007/s00520-019-04805-1

Grégoire, C., Bragard, I., Jerusalem, G., Etienne, A.-M., Coucke, P., Dupuis, G., et al. (2017). Group interventions to reduce emotional distress and fatigue in breast cancer patients: a 9-month follow-up pragmatic trial. *Br. J. Cancer* 117, 1442–1449. doi: 10.1038/bjc.2017.326

Grégoire, C., Faymonville, M.-E., Vanhaudenhuyse, A., Charland-Verville, V., Jerusalem, G., and Bragard, I. (2018a). Randomized controlled trial of an 8-week intervention combining self-care and hypnosis for post-treatment cancer patients: study protocol. *BMC Cancer* 18:1113. doi: 10.1186/s12885-018-5046-6

Grégoire, C., Faymonville, M.-E., Vanhaudenhuyse, A., Charland-Verville, V., Jerusalem, G., Willems, S., et al. (2020a). Effects of an intervention combining self-care and self-hypnosis on fatigue and associated symptoms in post-treatment cancer patients: a randomized-controlled trial. *Psycho-Oncology* 29, 1165–1173. doi: 10.1002/pon.5395

Grégoire, S., Gagnon, J., Lachance, L., Shankland, R., Dionne, F., Kotsou, I., et al. (2020b). Validation of the English and French versions of the multidimensional psychological flexibility inventory short form (MPFI-24). *J. Contextual Behav. Sci.* 18, 99–110. doi: 10.1016/j.jcbs.2020.06.004

Grégoire, C., Nicolas, H., Bragard, I., Delevallez, F., Merckaert, I., Razavi, D., et al. (2018b). Efficacy of a hypnosis-based intervention to improve well-being during cancer: a comparison between prostate and breast cancer patients. *BMC Cancer* 18:677. doi: 10.1186/s12885-018-4607-z

Grégoire, C., Sombrun, C., Gosseries, O., and Vanhaudenhuyse, A. (2021). *La transe cognitive auto-induite: caractéristiques et applications thérapeutiques potentielles*. HEGEL: Penser par Soi-Même.

Hariton, E., and Locascio, J. J. (2018). Randomised controlled trials—the gold standard for effectiveness research. *BJOG* 125:1716. doi: 10.1111/1471-0528.15199

Heck, M., Wagener, A., and Simon, J. (2021). The sense of presence: Validation of a new presence questionnaire. Available at: https://orbi.uliege.be/handle/2268/260787 (Accessed June 8, 2021).

Hoang, H. T. X., Molassiotis, A., Chan, C. W., Nguyen, T. H., and Liep Nguyen, V. (2020). New-onset insomnia among cancer patients undergoing chemotherapy: prevalence, risk factors, and its correlation with other symptoms. *Sleep Breath.* 24, 241–251. doi: 10.1007/s11325-019-01839-x

Horneber, M., Bueschel, G., Dennert, G., Less, D., Ritter, E., and Zwahlen, M. (2012). How many cancer patients use complementary and alternative medicine: a systematic review and metaanalysis. *Integr. Cancer Ther.* 11, 187–203. doi: 10.1177/1534735411423920

Hove, M. J., Stelzer, J., Nierhaus, T., Thiel, S. D., Gundlach, C., Margulies, D. S., et al. (2016). Brain network reconfiguration and perceptual decoupling during an absorptive state of consciousness. *Cereb. Cortex* 26, 3116–3124. doi: 10.1093/cercor/bhv137

Huels, E. R., Kim, H., Lee, U., Bel-Bahar, T., Colmenero, A. V., Nelson, A., et al. (2021). Neural correlates of the shamanic state of consciousness. *Front. Hum. Neurosci.* 15:610466. doi: 10.3389/fnhum.2021.610466

Jensen, M. P., Adachi, T., Tomé-Pires, C., Lee, J., Osman, Z. J., and Miró, J. (2015). Mechanisms of hypnosis: toward the development of a biopsychosocial model. *Int. J. Clin. Exp. Hypn.* 63, 34–75. doi: 10.1080/00207144.2014.961875

John, G. M., Hershman, D. L., Falci, L., Shi, Z., Tsai, W.-Y., and Greenlee, H. (2016). Complementary and alternative medicine use among US cancer survivors. *J. Cancer Surviv.* 10, 850–864. doi: 10.1007/s11764-016-0530-y

Johnson, M. K., Foley, M. A., Suengas, A. G., and Raye, C. L. (1988). Phenomenal characteristics of memories for perceived and imagined autobiographical events. *J. Exp. Psychol. Gen.* 117, 371–376. doi: 10.1037/0096-3445.117.4.371

Joly, F., Lange, M., Rigal, O., Correia, H., Giffard, B., Beaumont, J. L., et al. (2012). French version of the functional assessment of cancer therapy-cognitive function (FACT-cog) version 3. *Support Care Cancer* 20, 3297–3305. doi: 10.1007/s00520-012-1439-2

Joly, F., Rigal, O., Noal, S., and Giffard, B. (2011). Cognitive dysfunction and cancer: which consequences in terms of disease management? *Psychooncology* 20, 1251–1258. doi: 10.1002/pon.1903

Jones, J. M., Olson, K., Catton, P., Catton, C. N., Fleshner, N. E., Krzyzanowska, M. K., et al. (2016). Cancer-related fatigue and associated disability in post-treatment cancer survivors. *J. Cancer Surviv.* 10, 51–61. doi: 10.1007/s11764-015-0450-2

Kawai, N., Honda, M., Nishina, E., Yagi, R., and Oohashi, T. (2017). Electroencephalogram characteristics during possession trances in healthy individuals. *Neuroreport* 28, 949–955. doi: 10.1097/WNR.0000000000000857

Kılıç, A., Hudson, J., McCracken, L. M., Ruparelia, R., Fawson, S., and Hughes, L. D. (2021). A systematic review of the effectiveness of self-compassion-related interventions for individuals with chronic physical health conditions. *Behav. Ther.* 52, 607–625. doi: 10.1016/j.beth.2020.08.001

Kim, Y. H., Kim, H. J., Ahn, S. D., Seo, Y. J., and Kim, S. H. (2013). Effects of meditation on anxiety, depression, fatigue, and quality of life of women undergoing radiation therapy for breast cancer. *Complement. Ther. Med.* 21, 379–387. doi: 10.1016/j.ctim.2013.06.005

King, M., Nazareth, I., Lampe, F., Bower, P., Chandler, M., Morou, M., et al. (2005). Impact of participant and physician intervention preferences on randomized trials: a systematic review. *JAMA* 293, 1089–1099. doi: 10.1001/jama.293.9.1089

Koh, W.-J., Abu-Rustum, N. R., Bean, S., Bradley, K., Campos, S. M., Cho, K. R., et al. (2019). Cervical cancer, version 3.2019, NCCN clinical practice guidelines in oncology. *J. Natl. Compr. Cancer Netw.* 17, 64–84. doi: 10.6004/jnccn.2019.0001

Kotsou, I., and Heeren, A. (2011). Pleine conscience et acceptation: Les thérapies de la troisième vague. Available at: https://dial.uclouvain.be/pr/boreal/object/boreal:93652 (Accessed June 3, 2021).

Kotsou, I., and Leys, C. (2016). Self-compassion scale (SCS): psychometric properties of the French translation and its relations with psychological well-being, affect and depression. *PLoS One* 11:e0152880. doi: 10.1371/journal.pone.0152880

Kowalski, C. J., and Mrdjenovich, A. J. (2013). Patient preference clinical trials: why and when they will sometimes be preferred. *Perspect. Biol. Med.* 56, 18–35. doi: 10.1353/pbm.2013.0004

Kwak, A., Jacobs, J., Haggett, D., Jimenez, R., and Peppercorn, J. (2020). Evaluation and management of insomnia in women with breast cancer. *Breast Cancer Res. Treat.* 181, 269–277. doi: 10.1007/s10549-020-05635-0

Landry, M., Lifshitz, M., and Raz, A. (2017). Brain correlates of hypnosis: a systematic review and meta-analytic exploration. *Neurosci. Biobehav. Rev.* 81, 75–98. doi: 10.1016/j.neubiorev.2017.02.020

Lin, C., Clark, R., Tu, P., Bosworth, H. B., and Zullig, L. L. (2017). Breast cancer oral anti-cancer medication adherence: a systematic review of psychosocial motivators and barriers. *Breast Cancer Res. Treat.* 165, 247–260. doi: 10.1007/s10549-017-4317-2

Loh, K. P., Zittel, J., Kadambi, S., Pandya, C., Xu, H., Flannery, M., et al. (2018). Elucidating the associations between sleep disturbance and depression, fatigue, and pain in older adults with cancer. *J. Geriatr. Oncol.* 9, 464–468. doi: 10.1016/j.jgo.2018.02.006

Lomas, T. (2015). A systematic review of the neurophysiology of mindfulness on EEG oscillations. *Neurosci. Biobehav. Rev.* 57, 401–410. doi: 10.1016/j.neubiorev.2015.09.018

Ma, Y., He, B., Jiang, M., Yang, Y., Wang, C., Huang, C., et al. (2020). Prevalence and risk factors of cancer-related fatigue: a systematic review and meta-analysis. *Int. J. Nurs. Stud.* 111:103707. doi: 10.1016/j.ijnurstu.2020.103707

Mackinnon, C. (2012). *Shamanism and Spirituality in Therapeutic Practice: An Introduction*. London, UK: Singing Dragon.

Maclean, K. A., Leoutsakos, J.-M. S., Johnson, M. W., and Griffiths, R. R. (2012). Factor analysis of the mystical experience questionnaire: a study of experiences occasioned by the hallucinogen psilocybin. *J. Sci. Study Relig.* 51, 721–737. doi: 10.1111/j.1468-5906.2012.01685.x

Mainieri, A. G., Peres, J. F. P., Moreira-Almeida, A., Mathiak, K., Habel, U., and Kohn, N. (2017). Neural correlates of psychotic-like experiences during spiritual-trance state. *Psychiatry Res. Neuroimaging* 266, 101–107. doi: 10.1016/j.pscychresns.2017.06.006

Mehnert, A., Hartung, T. J., Friedrich, M., Vehling, S., Brähler, E., Härter, M., et al. (2018). One in two cancer patients is significantly distressed: prevalence and indicators of distress. *Psychooncology* 27, 75–82. doi: 10.1002/pon.4464

Merckelbach, H., Horselenberg, R., and Muris, P. (2001). The creative experiences questionnaire (CEQ): a brief self-report measure of fantasy proneness. *Personal. Individ. Differ.* 31, 987–995. doi: 10.1016/S0191-8869(00)00201-4

Mitchell, A. J., Chan, M., Bhatti, H., Halton, M., Grassi, L., Johansen, C., et al. (2011). Prevalence of depression, anxiety, and adjustment disorder in oncological, haematological, and palliative-care settings: a meta-analysis of 94 interview-based studies. *Lancet Oncol.* 12, 160–174. doi: 10.1016/S1470-2045(11)70002-X

Neff, K. D., and Germer, C. K. (2013). A pilot study and randomized controlled trial of the mindful self-compassion program. *J. Clin. Psychol.* 69, 28–44. doi: 10.1002/jclp.21923

Neff, K. D., Knox, M. C., Long, P., and Gregory, K. (2020). Caring for others without losing yourself: an adaptation of the mindful self-compassion program for healthcare communities. *J. Clin. Psychol.* 76, 1543–1562. doi: 10.1002/jclp.23007

Neufeld, N. J., Elnahal, S. M., and Alvarez, R. H. (2017). Cancer pain: a review of epidemiology, clinical quality and value impact. *Future Oncol.* 13, 833–841. doi: 10.2217/fon-2016-0423

Ngamkham, S., Holden, J. E., and Smith, E. L. (2019). A systematic review: mindfulness intervention for cancer-related pain. *Asia Pac. J. Oncol. Nurs.* 6, 161–169. doi: 10.4103/apjon.apjon_67_18

Oohashi, T., Kawai, N., Honda, M., Nakamura, S., Morimoto, M., Nishina, E., et al. (2002). Electroencephalographic measurement of possession trance in the field. *Clin. Neurophysiol.* 113, 435–445. doi: 10.1016/S1388-2457(02)00002-0

Ospina, M. B., Bond, K., Karkhaneh, M., Buscemi, N., Dryden, D. M., Barnes, V., et al. (2008). Clinical trials of meditation practices in health care: characteristics and quality. *J. Altern. Complement. Med.* 14, 1199–1213. doi: 10.1089/acm.2008.0307

Patsopoulos, N. A. (2011). A pragmatic view on pragmatic trials. *Dialogues Clin. Neurosci.* 13, 217–224. doi: 10.31887/DCNS.2011.13.2/npatsopoulos

Peres, J. F., Moreira-Almeida, A., Caixeta, L., Leao, F., and Newberg, A. (2012). Neuroimaging during trance state: a contribution to the study of dissociation. *PLoS One* 7:e49360. doi: 10.1371/journal.pone.0049360

Preference Collaborative Review Group (2008). Patients' preferences within randomised trials: systematic review and patient level meta-analysis. *BMJ* 337:a1864. doi: 10.1136/bmj.a1864

Pullens, M. J. J., De Vries, J., and Roukema, J. A. (2010). Subjective cognitive dysfunction in breast cancer patients: a systematic review. *Psychooncology* 19, 1127–1138. doi: 10.1002/pon.1673

Riba, M. B., Donovan, K. A., Andersen, B., Braun, I., Breitbart, W. S., Brewer, B. W., et al. (2019). Distress management, version 3.2019, NCCN clinical practice guidelines in oncology. *J. Natl. Compr. Cancer Netw.* 17, 1229–1249. doi: 10.6004/jnccn.2019.0048

Sanford, S. D., Beaumont, J. L., Butt, Z., Sweet, J. J., Cella, D., and Wagner, L. I. (2014). Prospective longitudinal evaluation of a symptom cluster in breast cancer. *J. Pain Symptom Manag.* 47, 721–730. doi: 10.1016/j.jpainsymman.2013.05.010

Santoso, A. M. M., Jansen, F., de Vries, R., Leemans, C. R., van Straten, A., and Verdonck-de Leeuw, I. M. (2019). Prevalence of sleep disturbances among head and neck cancer patients: a systematic review and meta-analysis. *Sleep Med. Rev.* 47, 62–73. doi: 10.1016/j.smrv.2019.06.003

Savard, M.-H., Savard, J., Simard, S., and Ivers, H. (2005). Empirical validation of the insomnia severity index in cancer patients. *Psychooncology* 14, 429–441. doi: 10.1002/pon.860

Sawni, A., and Breuner, C. C. (2017). Clinical hypnosis, an effective mind-body modality for adolescents with behavioral and physical complaints. *Children* 4:19. doi: 10.3390/children4040019

Schieber, K., Niecke, A., Geiser, F., Erim, Y., Bergelt, C., Büttner-Teleaga, A., et al. (2019). The course of cancer-related insomnia: don't expect it to disappear after cancer treatment. *Sleep Med.* 58, 107–113. doi: 10.1016/j.sleep.2019.02.018

Sedgwick, P. (2013). What is a patient preference trial? *BMJ* 347:f5970. doi: 10.1136/bmj.f5970

Sidky, H. (2009). A Shaman's cure: the relationship between altered states of consciousness and shamanic healing1. *Anthropol. Conscious.* 20, 171–197. doi: 10.1111/j.1556-3537.2009.01016.x

Smets, E. M., Garssen, B., Bonke, B., and De Haes, J. C. (1995). The multidimensional fatigue inventory (MFI) psychometric qualities of an instrument to assess fatigue. *J. Psychosom. Res.* 39, 315–325. doi: 10.1016/0022-3999(94)00125-O

Sombrun, C. (2006). *Mon Initiation Chez les Chamanes*. Paris: Pocket.

Sombrun, C. (2012). *Les Esprits de la steppe: Avec les derniers chamanes de Mongolie*. Paris: Albin Michel.

Spiegel, D. (1985). The use of hypnosis in controlling cancer pain. *CA Cancer J. Clin.* 35, 221–231. doi: 10.3322/canjclin.35.4.221

Spiegel, D. (1991). Neurophysiological correlates of hypnosis and dissociation. *J. Neuropsychiatr. Clin. Neurosci.* 3, 440–445. doi: 10.1176/jnp.3.4.440

Suh, H.-W., Jeong, H. Y., Hong, S., Kim, J. W., Yoon, S. W., Lee, J. Y., et al. (2021). The mindfulness-based stress reduction program for improving sleep quality in cancer survivors: a systematic review and meta-analysis. *Complement. Ther. Med.* 57:102667. doi: 10.1016/j.ctim.2021.102667

Sung, H., Ferlay, J., Siegel, R. L., Laversanne, M., Soerjomataram, I., Jemal, A., et al. (2021). Global cancer statistics 2020: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. *CA Cancer J. Clin.* 71, 209–249. doi: 10.3322/caac.21660

Thibaut, A., Panda, R., Annen, J., Sanz, L. R. D., Naccache, L., Martial, C., et al. (2021). Preservation of brain activity in unresponsive patients identifies MCS star. *Ann. Neurol.* 90, 89–100. doi: 10.1002/ana.26095

Vanhaudenhuyse, A., Laureys, S., and Faymonville, M.-E. (2014). Neurophysiology of hypnosis. *Neurophysiol. Clin.* 44, 343–353. doi: 10.1016/j.neucli.2013.09.006

Vanhaudenhuyse, A., Nyssen, A.-S., and Faymonville, M.-E. (2020). Recent insight on how the neuroscientific approach helps clinicians. *OBM Int. Complement. Med.* 5:20. doi: 10.21926/obm.icm.2002028

Watson, M., Greer, S., Young, J., Inayat, Q., Burgess, C., and Robertson, B. (1988). Development of a questionnaire measure of adjustment to cancer: the MAC scale. *Psychol. Med.* 18, 203–209. doi: 10.1017/S0033291700002026

Weitzenhoffer, A. M. (2002). Scales, scales and more scales. *Am. J. Clin. Hypn.* 44, 209–219. doi: 10.1080/00029157.2002.10403481

Zigmond, A. S., and Snaith, R. P. (1983). The hospital anxiety and depression scale. *Acta Psychiatr. Scand.* 67, 361–370. doi: 10.1111/j.1600-0447.1983.tb09716.x

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

frontiers in Psychology

HYPOTHESIS AND THEORY
published: 16 March 2022
doi: 10.3389/fpsyg.2022.804896

Is Higher-Order Misrepresentation Empirically Plausible? An Argument From Corruption

Asger Kirkeby-Hinrup[1,2]*

[1]*Theoretical Philosophy, Department of Philosophy and Cognitive Science, Lund University, Lund, Sweden,* [2]*Cognitive Neuroscience Research Unit, Center of Functionally Integrative Neuroscience, Aarhus University, Aarhus, Denmark*

OPEN ACCESS

Edited by:
Xerxes D. Arsiwalla,
Pompeu Fabra University,
Spain

Reviewed by:
Tomáš Marvan,
Academy of Sciences of the Czech Republic (ASCR), Czechia
Michał Klincewicz,
Tilburg University, Netherlands

****Correspondence:***
Asger Kirkeby-Hinrup
asger.kirkeby-hinrup@fil.lu.se

Specialty section:
This article was submitted to Consciousness Research, a section of the journal Frontiers in Psychology

Received: *29 October 2021*
Accepted: *23 February 2022*
Published: *16 March 2022*

Citation:
Kirkeby-Hinrup A (2022) Is Higher-Order Misrepresentation Empirically Plausible? An Argument From Corruption. Front. Psychol. 13:804896. doi: 10.3389/fpsyg.2022.804896

I present an empirically based argument for the plausibility of misrepresentation as posited by some higher-order theories of consciousness. The argument relies on the assumption that conscious states are generated by processes in the brain. The underlying idea is that if the brain generates conscious states then misrepresentation may occur. The reason for this is that brain states can be corrupted and, accordingly, a conscious state that is at least partly caused by a corrupted brain state may be a misrepresentation. Our body of knowledge from cognitive and behavioral neuroscience lends support to the idea that corruption of neural states is both possible and relatively frequent. If this is the case, I argue, it is plausible that occasionally such corruption may result in misrepresentation. I support this claim by arguing that the most prevalent theoretical alternative to the occurrence of misrepresentation—the so-called no-consciousness reply—seems less supported by our current knowledge in the domain of consciousness and cognition. This way of arguing for misrepresentation is different from other empirically based arguments in the debate because it is a meta-level argument resting on a general premise that most participants in the debate can accept.

Keywords: higher-order theory, misrepresentation, consciousness, HOT theory, functionalism, materialism, higher-order misrepresentation

INTRODUCTION

One of the central questions for a theory of consciousness is what accounts for the difference between the mental states that are conscious, and those that are not. One group of theories argues that what provides an individual with a conscious experience of a mental state *p1* is the presence of another mental state *p2* that has *p1* as its intentional object. Because *p2* is about another mental state, *p2* is considered a higher-order state. Therefore, theories that explain consciousness in terms of higher-order states are called higher-order theories.[1] Higher-order theories come in a wide variety of forms (Rosenthal, 1997; Carruthers, 2003; Lau, 2007; Gennaro, 2012; Coleman, 2015). One question that higher-order theories face is the distinctive roles played by *p1* and *p2*. For instance, is it possible that a higher-order state can misrepresent what mental state an individual is in? This is the question of misrepresentation. Those who endorse misrepresentation argue that the presence of *p2* is sufficient to generate a conscious experience of *p1* regardless of whether *p1* exists. Those who reject misrepresentation deny that a higher-order state—in itself—is sufficient for the individual to be in a conscious state.

In her seminal paper on misrepresentation, Neander (1998) distinguished between two types of misrepresentation. In *mild* misrepresentation, the higher-order state inaccurately represents the first-order state that it is about. Mild misrepresentation can be exemplified by a case where the individual has a first-order visual representation of *red* but the higher-order state makes the individual consciously experience *seeing blue*. In *radical* misrepresentation, the individual has a conscious experience of being in a state that she is not in. Radical misrepresentation occurs when a higher-order state exists but the lower-order state, that the higher-order state represents the individual as being in, does not. David Rosenthal, a proponent of misrepresentation, has criticized the distinction between mild and radical misrepresentation. Rosenthal argues that the line between mild and radical misrepresentation ultimately is arbitrary. Rosenthal (2004, p. 32) writes as:

"Suppose my higher-order awareness is of a state with the property P, but the target is not P, but rather Q. We could say that the higher-order awareness misrepresents the target, but we could equally well say that it's an awareness of a state that does not occur. The more dramatic the misrepresentation, the greater the temptation to say the target is absent."

So, Rosenthal argues that to ask whether a higher-order state is a mild or radical misrepresentation is misguided, because there is no non-arbitrary way to decide whether a higher-order thought (HOT) misrepresents its target or is about an absent target. How we describe the situation seems to be a matter of degree. If Rosenthal is right about this, it appears that the distinction between mild and radical misrepresentation collapses into radical misrepresentation.

Rosenthal holds that when misrepresentation occurs, the individual consciously experiences whatever the occurring higher-order thought represents her as experiencing. Rosenthal thinks that the possibility of discrepancies between a higher-order state and its target follows naturally from his theory, and is not only possible, but also fully coherent and theoretically harmless (Rosenthal, 2011, 2012). Many other proponents of HOT theory share this sentiment (e.g., Matey, 2006, 2011; Weisberg, 2010, 2011; Pereplyotchik, 2013; Berger, 2014).

In this paper, I present an empirically based argument for misrepresentation as posited by the higher-order theories of consciousness. This way of arguing for misrepresentation follows in the footsteps of earlier work in the debate by the advocates of misrepresentation and their opponents. For instance, Lau and Rosenthal (2011, p. 396) present empirical evidence from cognitive neuroscience, they argue provides the higher-order view with "substantial empirical plausibility." What is implied in this line of thinking is that experimental and clinical findings carry evidential weight in the theoretical domain. Similarly, Lau and Brown (2019) take aim specifically at the issue of misrepresentation and present empirical cases they interpret as the occurrence of misrepresentation. They argue that the fact that empirical cases exist shows that misrepresentation is not just a hypothetical conceptual problem, but that a successful theory will need to explain these cases. They conclude that in this respect the higher-order theory fares better than its competitors. Not only proponents of higher-order theory have argued on empirical grounds in the debate. In fact, leveraging empirical evidence is becoming increasingly prevalent in debates between competing theories of consciousness. For instance, Kozuch (2014, p. 722) acknowledges that one virtue of the higher-order theories is the amenability to empirical confirmation or disconfirmation. Kozuch proceeds to argue that evidence from lesions to the prefrontal cortex tells against the higher-order account. Similarly, other participants in the debate (e.g., Beeckmans, 2007; Malach, 2011; Sebastián, 2014) have leveraged empirically based arguments against the higher-order theories. Finally, specifically in relation to misrepresentation, principled (but not currently feasible) ways of testing for this have been suggested (Kirkeby-Hinrup, 2020). However, some empirically based arguments proposed in support of misrepresentation recently also have had objections leveraged against them (e.g., Kirkeby-Hinrup, 2014, 2016; Brinck and Kirkeby-Hinrup, 2017).

The assumption that empirical data may arbitrate between philosophical theories that are on equal footing on conceptual grounds, i.e., by providing a basis for an inference to the best explanation is gaining traction within current debates on consciousness, even in light of warnings about the work empirical evidence can do for us in this regard (Hohwy, 2009; Fink, 2016; Klein et al., 2020; Overgaard and Kirkeby-Hinrup, 2021). Be that as it may, the assumption that empirical evidence has an important role to play is shared by many philosophers of mind, who disagree on almost everything else. For instance, Josh Weisberg (2013) suggests that the right way to approach the study of consciousness is through empirical data. Similarly, Brown (2012) suggests that any theory of consciousness that is going to be physically realistic must take into account the nature of the brain and its states. The importance of empirical evidence in the debate is underscored by Block (2007, p. 486), when he suggests that "the familiar default 'method' of inference to the best explanation, that is, the approach of looking for the framework that makes the most sense of all the data […]" is the best way to examine the relation between phenomenal consciousness and brain states. Recently, steps have been taken to attempt carrying this out in practice

[1]There are of course alternatives to the higher-order thought theories of consciousness. Each of these accounts are characterized by the endorsement of an alternative mechanism through which mental states are rendered conscious. Some of these accounts, in particular first-order (or reflexive) theories (Kriegel, 2003a,b, 2007; Lamme, 2004), are held to preclude the possibility of misrepresentation. Others, such as workspace (Baars, 1996, 1997, 2005; Mashour et al., 2020) and integration (Tononi, 2005; Tononi et al., 2016) theories, have not—to the best of my knowledge—been explicitly considered in relation the misrepresentation debate.

(Kirkeby-Hinrup and Fazekas, 2021). In later work (Block, 2009, p. 1120), further notes that "it is hard to avoid the impression that the biology of the brain is what matters to consciousness—at least the kind we have."

In the rest of this paper, I will make the case that misrepresentation seems plausible given what we know about the brain. Unlike some previously proposed empirical arguments (e.g., Lau and Rosenthal, 2011; Lau and Brown, 2019), I do not take my starting point in neither a particular theory of consciousness, nor on a concrete empirical phenomenon. The argument I develop merely relies on the assumption that conscious states are generated in, and by, the brain. This assumption can be cashed out in different ways depending on how one conceives of the mind–brain relationship (see endnote iii). However, given that many theories of consciousness, and in particular most higher-order theories, are in the business of naturalizing the mind, they share this assumption. The underlying idea of the argument is that if the brain generates conscious states then misrepresentation may occur. The reason for this is that brain states can be corrupted and, accordingly, a conscious state that is at least partly caused by a corrupted brain state may be a misrepresentation. Call this the argument from corruption (AFC). The way the AFC argues for misrepresentation is different from other approaches to misrepresentation: the AFC is a meta-level argument resting on a general premise that most participants in the debate can accept (that conscious states somehow rely on brain activity). This means that the AFC does not take its starting point in a particular theory of consciousness but instead appeals to a view about the brain that is presumably shared by both proponents and opponents of misrepresentation. In support of this presumption, I gave examples above of a range of participants in the debate who appear to share this view.

THE ARGUMENT FROM CORRUPTION

The AFC turns on a central assumption of those who are engaged in the project of naturalizing the mind. The assumption is that conscious states are generated in the brain, and consciousness thus depends on the integrity of its neural underpinnings. Given that this assumption is shared by most who oppose and who endorse misrepresentation, the AFC can proceed from common ground, thereby increasing the chance of making progress in the debate.

From this starting point, the AFC proceeds with the following question: If we think that conscious states are generated in the brain and we know that the physical makeup of the brain is susceptible to corruption, then why could corruption of the physical makeup of the brain not result in misrepresentation? Now obviously, as it stands, this question puts too much of the burden of proof on opponents of misrepresentation. So, the plausibility of the AFC will rely on an explication of the way corruption may result in misrepresentation as envisioned by the proponents of higher-order theory.[2] Doing this will take a few steps, and as an initial move, it is useful to isolate the two premises that form the basis of the question.

1. Conscious states are generated in the brain.
2. Brain states are susceptible to corruption.

For the AFC to be plausible, it is necessary to justify each of these two premises. Regarding the first premise, this is usually taken for granted in the debates between competing (empirical) theories of consciousness. I will take this for granted here, but further support for this stance can be found in the introduction, as well as prominent publications (e.g., Doerig et al., 2020), and the whole debate about the localization of the neural correlates of consciousness (Lamme, 2003, 2004; Bor and Seth, 2012; Meuwese et al., 2013; Frässle et al., 2014; Kozuch, 2014; Boly et al., 2017; Odegaard et al., 2017; Michel and Morales, 2020).

In the next two sections, I will defend the second premise and provide two ways of conceiving of corruption at a general level (the philosophically inclined also may consult this[3] lengthy endnote).

[2]One may, of course, have separate reasons for rejecting higher-order theory, and while the argument presented here—if successful—deals with one objecting to higher-order theory, it does not tell against competing theories of consciousness.

[3]For the philosophically inclined, here are two specified versions of the AFC deploying different standpoints on the way conscious states depend on brain activity. The positions most prevalent in current debates on consciousness are functionalism and materialism. Let us begin by framing AFC in functionalist terms (AFC-F). *AFC-F*
P1) The neural matter of the brain is corruptible.
P2) When neural matter is corrupted, any functions that are instantiated in it may malfunction.
P3) Conscious states are functional states of the brain.
C) Conscious states may malfunction.
It appears that the AFC-F is valid as it is presented here. In the main text, I presented considerations in support of the first and second premises. The third premise follows from a functionalist view of the mind–brain relationship. Of course, the justification of the functionalist view may need independent motivation, but because we are taking it for granted here, P3 appears to be granted as well. Thus, it appears AFC is compatible with a functionalist view of the mind–brain relationship. Moving on to the version of AFC framed in materialist terms (AFC-M). Here is one way to do it:
AFC-M
P1) The neural matter of the brain is corruptible.
P2) When neural matter is corrupted, any states of the matter may be corrupted.
P3) Conscious states are (contingently) identical to brain states.
C) Conscious states may be corrupted.
Again, the argument appears valid. The first and second premises have been secured in above. However, one might object that on AFC-M the second premise appears too weak. The second premise appears too weak if one espouses a token-identity theory. Token identity entails that the relevant states under consideration in AFC-M are numerically identical to neural matter. If the mental states are numerically identical to neural matter, then it is not the case that the mental states *may* be corrupted. Rather, the objector will claim, any corruption of the neural matter entails corruption of states of the matter. I think this objection is valid. However, if the objection is correct, it only precipitates the strength of the AFC-M. For this reason, we can be satisfied with the weaker formulation of the second premise. The third premise merely reiterates the materialist standpoint. This might need independent motivation, but its validity is taken for granted here. The upshot of AFC-M is that it appears that corruption of the neural underpinnings of conscious states directly entails corruption of the mental states. Therefore, it appears as though the AFC has good footing if one adopts materialism.
Finally, one might notice that the conclusions of AFC-F and AFC-M differ. The functionalist version concludes that conscious states *may malfunction*, while the materialist version concludes that conscious states *may be corrupted*. Therefore, one might ask whether both conclusions entail the possibility of misrepresentation. However, the difference in wording merely reflects the terminology of the framework, and the difference in wording therefore is inconsequential. The terms "malfunction" and "corruption" describe the same underlying phenomenon, *viz.* that something has gone awry in the neural machinery, which in turn may or may not affect the conscious experience of the individual.

With the two premises established, the central part of the question remains as: Why would it be impossible for corruption of the brain to result in misrepresentation? Readers familiar with the philosophy of science will surely recognize the induction problem lurking in the background here. Furthermore, asking opponents of a view to prove a negative is not a viable option. For these reasons, the crux of the AFC consists in pointing to the fact that corruption of the neural underpinnings of consciousness often results in a wide variety of surprising and counterintuitive phenomena. Given the prevalence and variety of such phenomena, the question then becomes whether we have any *empirical* reason to think misrepresentation could not result from corruption. The reason I accentuate *empirical* here is that some opponents of misrepresentation have theoretical or conceptual reasons for rejecting misrepresentation. Later (in The No-Consciousness Reply), I will consider, and reject, one prominent such reason.

Are Brain States Corruptible?

In this section, I will briefly motivate the second premise of the AFC that brain states can be corrupted. Given that the topic of this text is the possibility of misrepresentation, the premise is in need of two major specifications. First, because misrepresentation requires that an individual is conscious of *something*, the kind of corruption that is relevant to the argument here cannot be such that it extinguishes consciousness. This means that cases of severe brain trauma that leave the individual unconscious, in a coma, or dead (although being clear cases of corruption of the physical makeup or function of the brain) cannot be invoked here. Second, the meaning of the term *corruption* must be clear. I here take corruption to be any kind of event in—or state of—the brain that results in abnormal processing, where this is defined in opposition to neurotypical subjects or processing. Findings in the fields of behavioral and cognitive neuroscience clearly support the possibility of corruption. Indeed, these fields are concerned with coupling observations of behavioral or cognitive performance with their neural underpinnings, and in many cases, behavioral and cognitive performance is abnormal (see, Gazzaniga et al., 2014 for an extensive review). Significant portions of the brain sciences take their starting points in examinations of various forms of abnormal cognitive or behavioral phenomena and investigate their neural causes. I take this fact to be sufficient to show that brain states are corruptible and sufficient to establish the second premise in general. To boot, many of the arguments leveraged in the debate between higher-order theories and the opponents rely on varieties of lesions or otherwise corrupted neural processing (e.g., the rare Charles Bonnet syndrome discussed in Lau and Rosenthal, 2011). I submit that this shows that the second premise is not controversial in the particular context of misrepresentation either. Nevertheless, it is conducive to understanding the AFC to get a more detailed view of the kind of corruption that may be relevant. Therefore, I will next present two types of corruption that appear relevant to the possibility of misrepresentation.

Two Types of Corruption

To determine whether corruption can lead to misrepresentation we need to have a firmer grip on the notion of corruption and how it may work. Below, I distinguish between two types of corruption, based on considerations of how the brain processes and transfers information. The first type of corruption relates to the *transfer* of information across topographically distinct areas of the brain. I will call it *Corruption in Information Transfer* (CIT). CIT can be divided into two types.

The first type of CIT can be called *external* CIT. External CIT suggests that when information is transferred between distinct faculties something may go awry. What awry means in this context is that the information carrying signal is degraded or otherwise distorted in a way that affects the information embedded therein (e.g., as a result of degraded myelin sheaths or through the application of TMS). Even on the micro scale (such as in the signals from one neural ensemble to another), the transfer of information involves signals traveling across actual physical distances. It is also possible to envision external CIT occurring at the macro level. For instance, visual signals travel from the retina through the optic nerve to the visual cortex and beyond through the ventral and dorsal streams. To illustrate external CIT, imagine a messenger traveling with a bag of letters from one town to another. At one point, part of the road has been flooded and the bag of letters becomes wet, causing the ink of the letters to smudge. When the messenger arrives with the letters, their content has literally changed (how the recipients of the letters interpret the corrupted content is a separate question). In this analogy, the road is the neural pathway across which information is transferred. The two towns are the faculties between which information is transferred, and the letters are the information.

The second type of CIT may occur when information is transferred within a given faculty. When CIT is occurring in the transfer *within* a faculty, I will call it *internal* CIT. Because we know that many faculties (e.g., the visual system) are distributed across distinct topographical locations of the brain, information is often transferred internally within a faculty as processing is carried out. For instance, if one conceives of the visual system as comprising a faculty, it appears reasonable to say that this faculty is topographically distributed. It is distributed because visual input is processed in more than one place (e.g., the striate cortex and prestriate cortex). Furthermore, it is fairly well established that visual information is (initially) transferred sequentially through distinct topographical locations. At each stage in the sequence, the input received is processed for particular properties. Thus, the processing of, for example, spatial frequency and motion are handled separately. One might object that we should view each of these stages in the sequence as faculties on their own, rather than grouping them together into a large visual faculty. However, this is not an argument against CIT since on this view each of the faculties that belong to the visual system will still be distributed across several neural ensembles and thus will be susceptible to internal CIT when information is transferred among them.

For a useful analogy to illustrate internal CIT, imagine that a large corporation has hired a consultant to produce a report on some important issue. Once the report is received, it is passed through various departments of the company; each department adds their perspective and comments on the issue in question. The financial department adds some figures and some calculations of expenses and expected revenue. The marketing department produces an appendix concerning user segments, merchandize, advertising platforms, and so forth. Once the report has passed through all the relevant departments of the corporation it reaches the boardroom, the members of which will take some appropriate action based on the report. One can imagine that, at some point in the process of being shipped from one department to the next, a couple of pages of the report containing crucial notes or calculations get lost or become damaged. The upshot is that when the report finally reaches the boardroom its contents have been corrupted and the considerations of the board will be different than they would have been if the report had been intact. Importantly, the corruption of the report occurs *between* the departments, in the transfer of information. In the analogy, the report from the consultant is the information input to the faculty, the different departments are the internal parts of the faculty that process the input, and the boardroom is the output function of the faculty.

One might object that this analogy is too simple. Perhaps one finds it implausible that such an illustration maps into very complex neural circuitry. Perhaps one would insist that, for this analogy to be a reasonable description of neural processing, more than one department should be working on the report simultaneously. However, imagining a more complex corporation, with several input/output sections, and parallel processing, only increases the number of paths across which information must be transferred. This means that the possibility of corruption during transfer of information actually may increase with the complexity of the corporation (faculty). It worth noting that that parallel processing also may guard against corruption by maintaining the information in separate processing streams, which may decrease the impact of corruption to one stream. However, it is not clear that this will preempt the issues raised here, since it raises questions relating to how to arbitrate between diverging streams that originally contained the same information, possible corruption to such an arbitration mechanism, and retains the issue for cases where multiple processing streams do not obtain.

The second type of corruption one may envision is in the *processing* of the information of a given faculty or neural ensemble. Call this *Process Corruption* (PC). A range of cortical areas appears to be highly specialized. An example of specialized areas could be those comprising the visual system, where for instance V4 handles specific properties of the visual signal such as spatial frequency and orientation. When positing the possibility of PC, one envisions that the procedural integrity of faculties or neural ensembles may be corrupted. The result is that the faculties process information in abnormal ways. For a useful analogy to illustrate PC, we can imagine the corporation described above. As before, the report represents the information being transferred through different departments of the corporation. However, in PC, the corruption does not occur in the transfer from one department to another. To illustrate PC, we instead imagine that one of the departments makes a critical mistake. For instance, the financial department might use an erroneous model to predict the development of the market, or simply mistype numbers in the budget. Importantly, it is the processing by a particular entity that corrupts the information and yields the abnormal output.

One might wonder whether the possibility of PC pertains only to the structural level (e.g., faculties) or whether it can occur at lower levels as well (e.g., neural ensembles). Let us consider the structural level first. In visual agnosia, individuals fail to process some specific feature of visual input owing to corruption of the relevant specialized faculty in their visual system. The behavioral evidence and subjective reports from patients in cases of visual agnosia clearly indicate that the relevant feature is not processed normally. In many cases, the behavioral evidence and subjective reports are corroborated by neural imaging showing abnormalities in the relevant faculty. From this, it appears there is reason to think that PC can occur at least at the structural level.

Does it occur at lower levels as well? In response to this question, there are at least two lines of reply. The first line of reply asks whether, when PC occurs at the structural level, it is always an entire structure that is corrupted, or only some part of it. It does not seem that we need to posit that the entire structure must be corrupted for it to yield abnormal processing. Rather, it appears that corruption of some (perhaps integral) part of the structure may be sufficient for the structure to yield abnormal processing. If this is the case, then it appears that we have obtained low-level PC for free, simply by showing that structural PC is possible.

The second line of reply consists in switching the burden of proof to those who might want to argue that low-level PC cannot occur. Why, one may ask, should we not believe that PC could occur at low levels of processing? It seems there are reasons to think that it can (e.g., the first line of reply, and possibly others such as the delicateness of biological matter), but no obvious reasons to think that it cannot.

I do not purport that the two types of corruption considered here are the only types of corruption that can occur. Corruption might occur in ways not considered here. The purpose of the examples given here is merely to describe two fairly basic and uncontroversial types of corruption.

THE NO-CONSCIOUSNESS REPLY

If the AFC is convincing, this means that misrepresentation is empirically plausible. The operative word here is "empirically." Conceptually, most agree that misrepresentation is possible, at least in so far as one endorses a representational theory of consciousness, given that a representational relation does not seem to entail the existence of what is represented (e.g., it is possible to represent the easter bunny). At the theoretical level however, several opponents of misrepresentation have denied that misrepresentation *in fact* obtains. Importantly, the motivations for this denial are theoretical rather than empirical. In this section, I will evaluate the so-called "no-consciousness reply" (Gennaro, 2004, 2006; Wilberg, 2010) given that this can be seen as reminiscent of an empirical claim. In brief, the no-consciousness reply accepts that occasionally a higher-order state may misrepresent its target first-order state but claim that in those cases no conscious event will follow, regardless of the cause of the misrepresentation. That is, if a higher-order state misrepresents its target state, the individual will not consciously experience being in the target state.

When applied to the AFC, the no-consciousness reply would amount to accepting the premises but rejecting the conclusion. In other words, proponents of the no-consciousness reply may accept that consciousness relies on processes in the brain (premise 1) and that brain processes are corruptible (premise 2) but reject that cases where misrepresentation occurs due to corruption can yield conscious states. One way to do this for proponents of the no-consciousness reply is to claim that precisely the neural processes underlying consciousness are *functionally fragile*, as it were. By "functionally fragile," one would mean that *any* corruption of the neural processes that generate consciousness would result in *no* conscious states being generated. Thus, the claim is that exactly the processes underlying consciousness are, in fact, *not* corruptible, but only can be destroyed. In its theoretical formulation, the no-consciousness reply amounts to a stipulation, for instance through the positing of a necessary intrinsic relation (in Gennaro's version between two proper parts of a complex mental state). Because this stipulation turns on an intuition that is not shared in the debate, its validity is problematic to assess and no consensus has emerged. Therefore, given that the claim is otherwise theoretically coherent and internally consistent, it appears the only way to evaluate objectively the functional fragility variant of the no-consciousness reply is to consider the empirical support for it.

There are neural processes and faculties that neuroscience suggests are empirically necessary for consciousness (Giacino et al., 2014). In addition to the necessary processes, there also are non-necessary processes involved in the production of consciousness at a given time. The non-necessary processes matter because in many cases these will modulate the *contents* of particular states of consciousness, even while they are not necessary for *being* conscious in the first place. Since misrepresentation is a matter of contents of states, what I here call non-necessary processes are highly relevant. For example, parts of the visual system may be damaged without neither consciousness, nor visual perception being extinguished completely, which goes to show that these processes cannot be *necessary* for consciousness and/or visual perception at the general level. Thus, non-necessary processes can be corrupted severely without extinguishing consciousness. This is supported by the fact that much of cognitive psychology and cognitive neuroscience is devoted exactly to investigating the symptoms of such corruption. An example of this is visual agnosia resulting from carbon monoxide poisoning (Gazzaniga et al., 2014, p. 225). The fact that certain processes involved in the generation of consciousness can be corrupted may suggest that some necessary processes might be corruptible as well. Inductive inference from the fact that many non-necessary processes are corruptible can be considered as support for this. Additionally, the fact that the necessary processes are instantiated in or identical to (see endnote *i* for this distinction) the same matter (*viz.* the brain) as the non-necessary processes lends some credibility to this inference.

Importantly, the AFC does not claim that corruption automatically generates misrepresentation. The claim here is not that any corruption automatically causes misrepresentation. If corruption is possible then it may destroy conscious states in some cases, just as is claimed by the proponents of the no-consciousness reply. In other cases, corruption may lead to degraded or otherwise flawed conscious states. Indeed, the list of the possible consequences of corruption may be very long. The AFC, I submit, is a reason to think that misrepresentation rightfully belongs on that list. The purpose here only is to make plausible that in some cases, corruption may result in misrepresentation. Given that there are at least inductive reasons based on the vast body of work in cognitive and behavioral neuroscience to think that corruption of both necessary and contingent neural processes may occur without extinguishing consciousness, the onus must be on the proponents of the no-consciousness reply to provide empirical support for their claim. Absent empirical reasons to think otherwise, claiming that exactly the processes underlying conscious states are functionally fragile appears *ad-hoc*.

CONCLUDING REMARKS

I have put forward the AFC to argue that misrepresentation is empirically plausible. The AFC suggests that if corruption of the neural underpinnings of the generation of conscious states is possible, then occurrences of misrepresentation are plausible. Upon considering whether the no-consciousness reply could be leveraged as an objection to the AFC, I concluded that there appears to be no empirically based reason to endorse it. On the contrary, there is some inductive empirical support for the idea that the neural underpinnings of consciousness *can* be corrupted. If corruption is possible, this is reason to think that occurrences of misrepresentation in fact obtain.

It is worth mentioning that in my treatment of the no-consciousness reply, I mainly considered a version of the AFC suggesting misrepresentation may occur as a result of corruption of the necessary processes for generating consciousness. In addition to this, there is a weaker version of AFC positing that misrepresentation may occur as a result of corruption in non-necessary processes. The idea behind this weaker claim is that errors in early processing in the non-necessary processes (e.g., submodules of visual system) may propagate upstream and ultimately yield misrepresentation once the resulting states become conscious. While the weaker claim is certainly interesting, the purpose of the present text has been merely to suggest the empirical viability of AFC based on corruption of core processes involved in consciousness. However, there is no provision in the debates that the occurrence of misrepresentation must be the "fault" of the HOT or the faculty that generates HOTs. What matters for misrepresentation is that a HOT renders an individual conscious of being in a state the individual is not in? For mild misrepresentation, it matters whether there is an "original" first-order state that is misrepresented in some way, but this criterion can still be satisfied by AFC. In the introduction, I showed that most participants in the misrepresentation debate agree on the two basic premises that consciousness relies on the brain and that empirical evidence is pertinent to philosophical debates on consciousness. In concordance with these views, it seems the AFC has a role to play in our understanding of misrepresentation.

Importantly, what the AFC seeks to establish is only that misrepresentation is *plausible*. This is enough to put pressure on proponents of the no-consciousness reply or theories who otherwise object to misrepresentation. Some proponents of misrepresentation additionally may endorse the stronger claim that the frequency of misrepresentation is higher than the occasional malfunction. How exactly an argument for this further claim might look is not my concern here. Nevertheless, initially establishing the empirical *plausibility* of misrepresentation is an important step along the way to constructing such an argument.

DATA AVAILABILITY STATEMENT

The original contributions presented in the study are included in the article/supplementary material, and further inquiries can be directed to the corresponding author.

AUTHOR CONTRIBUTIONS

The author confirms being the sole contributor of this work and has approved it for publication.

FUNDING

All funding for this paper under the post-doc award by the Swedish Research Council (Vetenskapsraadet) grant VR-2018-06595. Content of this paper was not influenced by the funding agency.

REFERENCES

Baars, B. J. (1996). Understanding subjectivity: global workspace theory and the resurrection of the observing self. *J. Conscious. Stud.* 3, 211–216.

Baars, B. J. (1997). In the theatre of consciousness. Global workspace theory, a rigorous scientific theory of consciousness. *J. Conscious. Stud.* 4, 292–309.

Baars, B. J. (2005). Global workspace theory of consciousness: toward a cognitive neuroscience of human experience. *Prog. Brain Res.* 150, 45–53. doi: 10.1016/S0079-6123(05)50004-9

Beeckmans, J. (2007). Can higher-order representation theories pass scientific muster? *J. Conscious. Stud.* 14, 90–111.

Berger, J. (2014). Consciousness is not a property of states: a reply to Wilberg. *Philos. Psychol.* 27, 829–842. doi: 10.1080/09515089.2013.771241

Block, N. (2007). Consciousness, accessibility, and the mesh between psychology and neuroscience. *Behav. Brain Sci.* 30:481. doi: 10.1017/S0140525X07002786

Block, N. (2009). *Comparing the Major Theories of Consciousness The Cognitive Neurosciences*. Cambridge, MA, US: Massachusetts Institute of Technology. 1111–1122.

Boly, M., Massimini, M., Tsuchiya, N., Postle, B. R., Koch, C., and Tononi, G. (2017). Are the neural correlates of consciousness in the front or in the back of the cerebral cortex? Clinical and neuroimaging evidence. *J. Neurosci.* 37, 9603–9613. doi: 10.1523/JNEUROSCI.3218-16.2017

Bor, D., and Seth, A. K. (2012). Consciousness and the prefrontal parietal network: insights from attention, working memory, and chunking. *Front. Psychol.* 3:63. doi: 10.3389/fpsyg.2012.00063

Brinck, I., and Kirkeby-Hinrup, A. (2017). Change blindness in higher-order thought: misrepresentation or good enough? *J. Conscious. Stud.* 24, 50–73.

Brown, R. (2012). "The brain and its states," in *Being in Time: Dynamical Models of Phenomenal Experience*. eds. E. Shimon, F. Tomer and Z. Neta (Netherlands: John Benjamins Publishing).

Carruthers, P. (2003). *Phenomenal Consciousness: A `Naturalistic Theory*. London: Cambridge University Press.

Coleman, S. (2015). Quotational higher-order thought theory. *Philos. Stud.* 172, 2705–2733. doi: 10.1007/s11098-015-0441-1

Doerig, A., Schurger, A., and Herzog, M. H. (2020). Hard criteria for empirical theories of consciousness. *Cogn. Neurosci.* 12, 41–62. doi: 10.1080/17588928.2020.1772214

Fink, S. B. (2016). A deeper look at the "neural correlate of consciousness." *Front. Psychol.* 7:1044. doi: 10.3389/fpsyg.2016.01044

Frässle, S., Sommer, J., Jansen, A., Naber, M., and Einhäuser, W. (2014). Binocular rivalry: frontal activity relates to introspection and action but not to perception. *J. Neurosci.* 34, 1738–1747. doi: 10.1523/JNEUROSCI.4403-13.2014

Gazzaniga, M. S., Ivry, R. B., and Mangun, G. R. (2014). *Cognitive Neuroscience: The Biology of the Mind*. *4th Edn.* ed. W. W. Norton (London, UK).

Gennaro, R. J. (ed.) (2004). "Higher-order thoughts, animal consciousness, and misrepresentation," in *Higher-Order Theories of Consciousness: An Anthology* (Netherland: John Benjamins), 45–68.

Gennaro, R. J. (2006). "Between pure self-referentialism and the (extrinsic) HOT theory of consciousness," in *Self-Representational Approaches to Consciousness*. U. Kriegel, K. Williford (England: MIT Press.

Gennaro, R. J. (2012). *The Consciousness Paradox: Consciousness, Concepts, and Higher-Order Thoughts*. England: MIT Press.

Giacino, J. T., Fins, J. J., Laureys, S., and Schiff, N. D. (2014). Disorders of consciousness after acquired brain injury: the state of the science. *Nat. Rev. Neurol.* 10, 99–114. doi: 10.1038/nrneurol.2013.279

Hohwy, J. (2009). The neural correlates of consciousness: new experimental approaches needed? *Conscious. Cogn.* 18, 428–438. doi: 10.1016/j.concog.2009.02.006

Kirkeby-Hinrup, A. (2014). Why the rare Charles bonnet cases are not evidence of misrepresentation. *J. Philos. Res.* 39, 301–308. doi: 10.5840/jpr20148420

Kirkeby-Hinrup, A. (2016). Change blindness and misrepresentation. *Disputatio* 8, 37–56. doi: 10.2478/disp-2016-0002

Kirkeby-Hinrup, A. (2020). "A higher-order faculty and beyond." in *Beyond the Neural Correlates of Consciousness*. eds. M. Overgaard, J. Mogensen and A. Kirkeby-Hinrup (New York: Routledge), 131–152.

Kirkeby-Hinrup, A., and Fazekas, P. (2021). Consciousness and inference to the best explanation: compiling empirical evidence supporting the access-phenomenal distinction and the overflow hypothesis. *Conscious. Cogn.* 94:103173. doi: 10.1016/j.concog.2021.103173

Klein, C., Hohwy, J., and Bayne, T. (2020). Explanation in the science of consciousness: From the neural correlates of consciousness (NCCs) to the difference makers of consciousness (DMCs). *Philo. Mind Sci.* 1:60. doi: 10.33735/phimisci.2020.II.60

Kozuch, B. (2014). Prefrontal lesion evidence against higher-order theories of consciousness. *Philos. Stud.* 167, 721–746. doi: 10.1007/s11098-013-0123-9

Kriegel, U. (2003a). Consciousness as intransitive self-consciousness: two views and an argument. *Can. J. Philos.* 33, 103–132. doi: 10.1080/00455091.2003.10716537

Kriegel, U. (2003b). Consciousness, higher-order content, and the individuation of vehicles. *Synthese* 134, 477–504. doi: 10.1023/A:1022913321854

Kriegel, U. (2007). The same-order monitoring theory of consciousness. Synthesis. *Philosophica* 22, 361–384.

Lamme, V. A. F. (2003). Why visual attention and awareness are different. *Trends Cogn. Sci.* 7, 12–18. doi: 10.1016/S1364-6613(02)00013-X

Lamme, V. A. F. (2004). Separate neural definitions of visual consciousness and visual attention; a case for phenomenal awareness. *Neural Netw.* 17, 861–872. doi: 10.1016/j.neunet.2004.02.005

Lau, H. (2007). A higher order Bayesian decision theory of consciousness. *Prog. Brain Res.* 168, 35–48. doi: 10.1016/S0079-6123(07)68004-2

Lau, H., and Brown, R. (2019). "The emperor's new phenomenology? The empirical case for conscious experiences without first-order representations," in *Blockheads! Essays on Ned Block's Philosophy of Mind and Consciousness*. eds. A. Pautz and D. Stoljar (MIT Press), 171–197.

Lau, H., and Rosenthal, D. M. (2011). Empirical support for higher-order theories of conscious awareness. *Trends Cogn. Sci.* 15, 365–373. doi: 10.1016/j.tics.2011.05.009

Malach, R. (2011). Conscious perception and the frontal lobes: comment on Lau and Rosenthal. *Trends Cogn. Sci.* 15, 508–509. doi: 10.1016/j.tics.2011.09.001

Mashour, G. A., Roelfsema, P., Changeux, J.-P., and Dehaene, S. (2020). Conscious processing and the global neuronal workspace hypothesis. *Neuron* 105, 776–798. doi: 10.1016/j.neuron.2020.01.026

Matey, J. (2006). Two HOTs to handle: the concept of state consciousness in the higher-order thought theory of consciousness. *Philos. Psychol.* 19, 151–175. doi: 10.1080/09515080600554753

Matey, J. (2011). Reduction and the determination of phenomenal character. *Philos. Psychol.* 24, 291–316. doi: 10.1080/09515089.2011.556608

Meuwese, J. D., Post, R. A., Scholte, H. S., and Lamme, V. A. (2013). Does perceptual learning require consciousness or attention? *J. Cogn. Neurosci.* 25, 1579–1596. doi: 10.1162/jocn_a_00424

Michel, M., and Morales, J. (2020). Minority reports: consciousness and the prefrontal cortex. *Mind Lang.* 35, 493–513. doi: 10.1111/mila.12264

Neander, K. (1998). The division of phenomenal labor: a problem for representational theories of consciousness. *Nous* 32, 411–434. doi: 10.1111/0029-4624.32.s12.18

Odegaard, B., Knight, R. T., and Lau, H. (2017). Should a few null findings falsify prefrontal theories of conscious perception? *J. Neurosci.* 37, 9593–9602. doi: 10.1523/JNEUROSCI.3217-16.2017

Overgaard, M., and Kirkeby-Hinrup, A. (2021). Finding the neural correlates of consciousness will not solve all our problems. *Philo. Mind Sci.* 2:37. doi: 10.33735/phimisci.2021.37

Pereplyotchik, D. (2013). Some HOT family disputes: a critical review of The consciousness paradox by Rocco Gennaro. *Philos. Psychol.* 15:1. doi: 10.1080/09515089.2013.838817

Rosenthal, D. M. (1997). "A theory of consciousness," in *The Nature of Consciousness: Philosophical Debates*. eds. N. Block, O. Flanagan and G. Güzeldere (England: MIT Press), 729–753.

Rosenthal, D. M. (2004). "Varieties of higher-order theory," in *Higher-Order Theories of Consciousness*. ed. R. J. Gennaro (Netherlands: John Benjamins), 19–44.

Rosenthal, D. M. (2011). Exaggerated reports: reply to Block. *Analysis* 71, 431–437. doi: 10.1093/analys/anr039

Rosenthal, D. M. (2012). Higher-order awareness, misrepresentation and function. *Philo. Trans. Royal Soc. B* 367, 1424–1438. doi: 10.1098/rstb.2011.0353

Sebastián, M. Á. (2014). "Not a HOT dream," in *Consciousness Inside and Out: Phenomenology, Neuroscience, and the Nature of Experience*. *Vol. 6.* ed. R. Brown (Netherlands: Springer), 415–432.

Tononi, G. (2005). "Consciousness, information integration, and the brain," in *Progress in Brain Research. Vol. 50.* ed. L. Steven (United Kingdome: Elsevier), 109–126.

Tononi, G., Boly, M., Massimini, M., and Koch, C. (2016). Integrated information theory: from consciousness to its physical substrate. *Nat. Rev. Neurosci.* 17, 450–461. doi: 10.1038/nrn.2016.44

Weisberg, J. (2010). Misrepresenting consciousness. *Philos. Stud.* 154, 409–433. doi: 10.1007/s11098-010-9567-3

Weisberg, J. (2011). Abusing the notion of what-it's-like-ness: a response to Block. *Analysis* 71, 438–443. doi: 10.1093/analys/anr040

Weisberg, J. (2013). "Sweet dreams are made of This? A HOT response to Sebastián," in *Consciousness Inside and Out: Phenomenology, Neuroscience, and the Nature of Experience. Vol. 6.* ed. R. Brown (Netherlands: Springer), 433–443.

Wilberg, J. (2010). Consciousness and false HOTs. *Philos. Psychol.* 23, 617–638. doi: 10.1080/09515089.2010.514567

Conflict of Interest: The author declares that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

frontiers | Frontiers in Psychology

HYPOTHESIS AND THEORY
published: 12 May 2022
doi: 10.3389/fpsyg.2021.704270

A Relativistic Theory of Consciousness

Nir Lahav[1]* **and Zachariah A. Neemeh**[2,3]

[1] Department of Physics, Bar-Ilan University, Ramat Gan, Israel, [2] Department of Philosophy, The University of Memphis, Memphis, TN, United States, [3] Institute for Intelligent Systems, The University of Memphis, Memphis, TN, United States

In recent decades, the scientific study of consciousness has significantly increased our understanding of this elusive phenomenon. Yet, despite critical development in our understanding of the functional side of consciousness, we still lack a fundamental theory regarding its phenomenal aspect. There is an "explanatory gap" between our scientific knowledge of functional consciousness and its "subjective," phenomenal aspects, referred to as the "hard problem" of consciousness. The phenomenal aspect of consciousness is the first-person answer to "what it's like" question, and it has thus far proved recalcitrant to direct scientific investigation. Naturalistic dualists argue that it is composed of a primitive, private, non-reductive element of reality that is independent from the functional and physical aspects of consciousness. Illusionists, on the other hand, argue that it is merely a cognitive illusion, and that all that exists are ultimately physical, non-phenomenal properties. We contend that both the dualist and illusionist positions are flawed because they tacitly assume consciousness to be an absolute property that doesn't depend on the observer. We develop a conceptual and a mathematical argument for a relativistic theory of consciousness in which a system either has or doesn't have phenomenal consciousness *with respect to some observer.* Phenomenal consciousness is neither private nor delusional, just relativistic. In the frame of reference of the cognitive system, it will be observable (first-person perspective) and in other frame of reference it will not (third-person perspective). These two cognitive frames of reference are both correct, just as in the case of an observer that claims to be at rest while another will claim that the observer has constant velocity. Given that consciousness is a relativistic phenomenon, neither observer position can be privileged, as they both describe the same underlying reality. Based on relativistic phenomena in physics we developed a mathematical formalization for consciousness which bridges the explanatory gap and dissolves the hard problem. Given that the first-person cognitive frame of reference also offers legitimate observations on consciousness, we conclude by arguing that philosophers can usefully contribute to the science of consciousness by collaborating with neuroscientists to explore the neural basis of phenomenal structures.

Keywords: consciousness, phenomenology, qualia, relativity (physics), the hard problem of consciousness, mathematical formulization of consciousness

OPEN ACCESS

Edited by:
Peter beim Graben, Humboldt University of Berlin, Germany

Reviewed by:
Tomáš Marvan, Academy of Sciences of the Czech Republic (ASCR), Czechia
Mark Leonard Solms, University of Cape Town, South Africa

***Correspondence:**
Nir Lahav
freenl@gmail.com

Specialty section:
This article was submitted to Consciousness Research, a section of the journal Frontiers in Psychology

Received: *02 May 2021*
Accepted: *01 December 2021*
Published: *12 May 2022*

Citation:
Lahav N and Neemeh ZA (2022) A Relativistic Theory of Consciousness. Front. Psychol. 12:704270. doi: 10.3389/fpsyg.2021.704270

INTRODUCTION

The Hard Problem of Consciousness

As one of the most complex structures we know of nature, the brain poses a great challenge to us in understanding how higher functions like perception, cognition, and the self arise from it. One of its most baffling abilities is its capacity for conscious experience (van Gulick, 2014). Thomas Nagel (1974) suggests a now widely accepted definition of consciousness: a being is conscious just if there is "something that it is like" to be that creature, i.e., some subjective way the world seems or appears from the creature's point of view. For example, if bats are conscious, that means there is something it is like for a bat to experience its world through its echolocational senses. On the other hand, under deep sleep (with no dreams) humans are unconscious because there is nothing it is like for humans to experience their world in that state.

In the last several decades, consciousness has transformed from an elusive metaphysical problem into an empirical research topic. Nevertheless, it remains a puzzling and thorny issue for science. At the heart of the problem lies the question of the brute phenomena that we experience from a first-person perspective—e.g., what it is like to feel redness, happiness, or a thought. These qualitative states, or qualia, compose much of the phenomenal side of consciousness. These qualia are arranged into spatial and temporal patterns and formal structures in phenomenal experience, called eidetic or transcendental structures[1]. For example, while qualia pick out how a specific note sounds, eidetic structures refer to the temporal form of the whole melody. Hence, our inventory of the elusive properties of phenomenal consciousness includes both qualia and eidetic structures.

One of the central aspects of phenomenal features is privacy. It seems that my first-person feeling of happiness, for example, cannot be measured from any other third-person perspective. One can take indirect measurements of my heart rate or even measure the activity in the brain networks that create the representation of happiness in my mind, but these are just markers of the feeling that I have, and not the feeling itself (Block, 1995). From my first-person perspective, I don't feel the representation of happiness. I just feel happiness. Philosophers refer to this feature of phenomenal consciousness as "transparency": we seem to directly perceive things, rather than mental representations, even though mental representations mediate experience. But this feeling cannot be directly measured from the third-person perspective, and so is excluded from scientific inquiry. Yet, let us say that we identify a certain mental representation in a subject's brain as "happiness." What justifies us in calling it "happiness" as opposed to "sadness?" Perhaps it is correlated with certain physiological and behavioral measures. But, ultimately, the buck stops somewhere: at a subject's phenomenological report otherwise, we cannot know what this representation represents (Gallagher and Zahavi, 2020). If we categorize the representation as "happiness" but the subject insists they are sad, we are probably mistaken, not the subject. Phenomenal properties are seemingly private and, by some accounts, even beyond any physical explanation. For being physical means being public and measurable and is epistemically inconsistent with privacy. Take an electron, for example. We know it is physical because we can measure it with e.g., a cathode ray tube. If electrons were something that only I could perceive (i.e., if they were private), then we would not include them as physical parts of the scientific worldview. Affective scientists, for example, measure aspects of feeling all the time, like valence and arousal, but measuring these is not the same as measuring the feeling of happiness itself. What affective scientists measure are outcomes of the feeling, while the feeling itself is private and not measurable by the scientists. To date, it is not clear how to bridge the "explanatory gap" (Levine, 1983) between "private" phenomenal features and public, measurable features (e.g., neurocomputational structures), leaving us stuck in the "hard problem" of consciousness (Chalmers, 1996). This gap casts a shadow on the possibility for neuroscience to solve the hard problem because all explanations will always remain within a third-person perspective (e.g., neuronal firing patterns and representations), leaving the first-person perspective out of the reach of neuroscience. This situation divides consciousness into two separate aspects, the functional aspect and the phenomenal one (Block, 1995). The functional aspect ('functional consciousness'), is the objectively observable aspect of consciousness (Franklin et al., 2016; Kanai et al., 2019). [In that sense, it's similar to Ned Block's definition of access conscious, but with less constraints. All phenomenal consciousness has a functional aspect and vice versa, whereas for Block this strict equivalence doesn't hold (Block, 2011).] But the subjective aspect (phenomenal consciousness) is not directly observable except on the part of the person experiencing that conscious state. As we saw above, because they are private, phenomenal properties are distinct from any cognitive and functional property (which can be publicly observable). Any theory of consciousness should explain how to bridge this gap—How can functional, public, aspects give rise to phenomenal, private aspects?

Nevertheless, in recent decades, consciousness has become increasingly amenable to empirical investigation by focusing on its functional aspect, finally enabling us to begin to understand this enigmatic phenomenon. For example, we now have good evidence that consciousness doesn't occur in a single brain area. Rather, it seems to be a global phenomenon in widespread areas of the brain (Baars, 1988; Varela et al., 2001; Mashour et al., 2020).

In studying *functional consciousness*, we take consciousness to be a form of information processing and manipulation of representations, and we trace its functional or causal role within the cognitive system. Widely successful theories, such as global workspace theory (Baars, 1988; Dehaene, 2014; Franklin et al., 2016), attention schema theory (Graziano, 2019), recurrent processing theory (Lamme, 2006; Fahrenfort et al., 2007), and integrated information theory (Tononi, 2008; Oizumi et al., 2014) are virtually premised on this information processing account. Despite our advances in the study of functional consciousness, we still lack a convincing way to bridge the explanatory gap to phenomenal consciousness. The questions, "Why does it feel

[1]The terms are nearly synonymous, but 'transcendental' means additionally that one has performed the *epochç*. The *epochç* refers to a theoretical bracketing of the ontological status of phenomena—i.e., bracketing the existence of the world. The goal of the *epochç* is to not confuse external reality with the immanent features of phenomenal consciousness. Since this is not a concern of ours here, we stick with the term 'eidetic.' See Husserl (1925/2012).

like anything at all to process information?", and "How can this feeling be private?" still remain controversial. The hard problem primarily remains a philosophical rather than scientific question. Phenomenal consciousness must be the ultimate reference point for any scientific theory of consciousness. Ultimately, theories of consciousness as information processing, i.e., theories of functional consciousness, only approximate full-blown consciousness by abstracting away its phenomenal features. But they must ultimately refer to phenomenal features in order to give a full explanation of consciousness. Otherwise, there are no grounds for labeling them theories of consciousness, rather than theories of cognition or global informational access.

Let's take Integrated Information Theory (IIT), for example. IIT claims to bridge between the phenomenal and functional aspects of consciousness and to answer the question of what qualia are (Tononi et al., 2016). According to IIT, consciousness is the result of highly integrated information in the brain [which is "the amount of information generated by a complex of elements, above and beyond the information generated by its parts" (Tononi, 2008)]. A mathematical formula can determine how much integrated information, and thus how much consciousness, is present in any specific system. This measure is referred to as ϕ, and if ϕ is higher than zero, the system is conscious. In this situation, the system has a "maximally irreducible conceptual structure" which is identical to its phenomenal experience (Oizumi et al., 2014). These maximally irreducible conceptual structures are composed of integrated information states, and thus a quale is identical to a specific relation between integrated information states. The problem is that these maximally irreducible conceptual structures are physical states and even though they can be very complex, they are still public and observable from a third-person perspective, while qualia are private and only perceptible from the first-person perspective. How can there be an identity between opposite properties like public and privacy properties? Yet, this is what IIT suggests without any explanation about how to bridge this contradiction. Let's choose a conceptual structure such that it is a quale of happiness according to IIT. However, it cannot truly be the quale of happiness because, in principle, when we measure it, we observe a physical process and not the qualitative happiness itself (that's why it's not enough to measure the physical process, and eventually we need to ask the participants what they feel). Again, we remain stuck within the explanatory gap and the hard problem. In other words, IIT hasn't solved the hard problem and hasn't bridged the explanatory gap between private and public properties (Mindt, 2017). In order to avoid this problem, IIT needs to assume phenomenal consciousness as a primitive element separated from other physical elements like space and time. In that case, there is no gap to bridge because now these maximally irreducible conceptual structures are not qualia anymore. They are just physical structures that have properties that correspond to the properties of phenomenal consciousness, but that are not identical to them. In that sense, IIT just shows correspondence between two separate kinds of elements, phenomenal and physical. In the next paragraph we will see that this kind of solution is exactly what Chalmers refers to as 'naturalistic dualism.' The problem with that is that instead of understanding what phenomenal consciousness is and what the relations are between the physical world and the phenomenal world, it just assumes that phenomenal consciousness exists as a separate basic element in the world, and it does not solve how there can be any interaction between physical and phenomenal elements (Carroll, 2021).

This is not an issue of only IIT, but a problem to all physicalist's theories. As David Chalmers (1995) has put it, "the structure and dynamics of physical processes yield only more structure and dynamics, so structures and functions are all we can expect these processes to explain." The structure and dynamics of information tell us nothing of the story of how one gets from these public structures and dynamics to our private phenomenal experience (and how there can be an equality between these opposite properties). This is known as the 'structure and dynamics argument' (Alter, 2016; Chalmers, 2003; Mindt, 2017), stating that structure and dynamics alone are not enough to account for consciousness. This raises a concern about whether any physicalist theory can solve the hard problem of consciousness. In the next sections, however, we will show that physicalism is broader than describing only structures and dynamics, and that we can use this fact in order to solve the hard problem.

Views about the relation between phenomenal and functional consciousness exist across a spectrum. On one end, illusionists seek to erase the hard problem by referring only to functional consciousness, taking phenomenal properties to be cognitive illusions. On the other end, naturalistic dualists or panpsychists seek to promote phenomenal consciousness as a fundamental, non-material, and undecomposable constituent of nature (Chalmers, 2017). The controversy over phenomenal consciousness can be traced to one central problem: naturalization. The project of naturalization involves taking folk psychological concepts and subjecting them to physical laws and empirical scrutiny (Hutto, 2007). Illusionists take the current scientific approach to consciousness and argue that this eliminates the messy problem with supposedly private, immaterial qualia. According to them, functional consciousness generates an illusion of special phenomenal properties, which create the persistent "user illusion" (Dennett, 1991) of a first-person perspective. Dualists start from the same problem of naturalization, but take it that phenomenal consciousness is simply not amenable to third-person scientific inquiry due to its sui generis properties. What is needed, according to the naturalistic dualist, is an expanded understanding of what counts as "natural." "Given that reductive explanation fails, *nonreductive* explanation is the natural choice" (Chalmers, 2017, p. 359). Chalmers proposes that consciousness is a fundamental property, something like the strong nuclear force that is irreducible to other forces. A complete ontology of the natural world simply must include phenomenal consciousness as a basic, undecomposable constituent.

The Zombie Argument and the Paradox of Phenomenal Judgment

Chalmers (1996) discusses the logical possibility of a *zombie*, a being physically, cognitively, and behaviorally identical to a human, but lacking phenomenal consciousness altogether. One would think that such a creature would be dull, like a robot with basic automated responses, but this is not the case. Because a zombie is physically identical to a human, it means that it has the same cognitive system as us: a system that gets inputs from the environment, processes them, and creates behavior responses. In fact, there are no differences between the human and the zombie's cognitive dynamics, representations, and responses. But, for the zombie, there is nothing that is like to do all these processes. According to Chalmers (1996), the zombie has phenomenal judgments. This concept is very important for our argument, so let's examine it a bit. Phenomenal judgements are higher-order cognitive functions that humans and zombies have in common. Humans are aware of their experience and its contents and they can form judgments about it (e.g., when we think 'There is something red'), then, usually, they are led to make claims about it. These various judgments in the vicinity of consciousness are phenomenal judgments. They are not phenomenal states themselves, but they are about phenomenology. Phenomenal judgments are often reflected in claims and reports about consciousness, but they start as a mental process. Phenomenal judgments are themselves cognitive acts that can be explained by functional aspects like the manipulation of mental representations. That's why zombies also have phenomenal judgements. We can think of a judgment as what is left of a belief after any associated phenomenal property is subtracted. As a result, phenomenal judgements are part of the functional aspect of consciousness. As Chalmers puts it (1996, p.174):

> "Judgments can be understood as what I and my zombie twin have in common. A zombie does not have any conscious experience, but he claims that he does. My zombie twin judges that he has conscious experience, and his judgments in this vicinity correspond one-to-one to mine. He will have the same form, and he will function in the same way in directing behavior as mine... Alongside every conscious experience there is a content-bearing cognitive state. It is roughly information that is accessible to the cognitive system, available for verbal report, and so on."

In other words, phenomenal judgments can be described as the representations of a cognitive system that bear content about phenomenology (which are not necessarily linguistic representations. E.g., such representation can be a representation of the color of the apple). In this paper, we will identify functional consciousness with the creation of phenomenal judgments.

As a result of the zombie's capacity to create phenomenal judgements, we reach a peculiar situation: The zombie has functional consciousness, i.e., all the physical and functional conscious processes studied by scientists, such as global informational access. But there would be nothing it is like to have that global informational access and to be that zombie. All that the zombie cognitive system requires is the capacity to produce phenomenal judgments that it can later report. For example, if you asked it if it sees a red rose in front of it, using information processing, it might respond, "Yes, I'm definitely conscious of seeing a red rose," although it is ultimately mistaken and there is truly nothing that is like for the zombie to see that rose. In order to produce this phenomenal judgment, despite having no phenomenal consciousness, the zombie cognitive system needs representations and a central system with direct access to important information enabling it to generate behavioral responses. It needs direct access to perceptual information, a concept of self to differentiate itself from the world, an ability to access its own cognitive contents, and the capacity to reflect. Such a cognitive system could presumably reason about its own perceptions. It would report that it sees the red rose, and that it has some property over and above its structural and functional properties—phenomenal consciousness. Of course, this report would be mistaken. It is a paradoxical situation in which functional consciousness creates phenomenal judgments without the intervention of phenomenal consciousness—yet phenomenal judgments are purportedly about phenomenal consciousness. This paradox of phenomenal judgment (Chalmers, 1996) arises because of the independence of phenomenal consciousness from physical processes. The hidden assumption here is that consciousness is private. Consequently, it is not possible to measure it. It seems that one aspect of consciousness (physical, functional consciousness) can come without the other (phenomenal consciousness). For example, in IIT there is a variant of the phenomenal judgement paradox. According to IIT, there can be a cognitive system that will manipulate information and infer that it has phenomenal experience, that there is something that is like to be that cognitive system, yet it doesn't have any consciousness because its neural network creates all these judgements in a feed-forward way, meaning that $\phi = 0$ in the system (Doerig et al., 2019). Again, we see that even in IIT the functional part of consciousness (phenomenal judgements) can come without the phenomenal aspect of consciousness. This paradox arises because in IIT consciousness and cognitive content are not conditioned to correspond to each other. As a result, although the cognitive system has phenomenal judgements about phenomenal consciousness, still, it is just a zombie with no phenomenal consciousness because $\phi = 0$ in the system (this kind of a zombie is termed functional zombie, see Oizumi et al., 2014; Tononi et al., 2016).

In order to solve this paradox, we need to explain two aspects of consciousness: How there could be natural phenomena that are private and thus independent of physical processes (or how come they *seem* private), and what the exact relationship between cognitive content and phenomenal consciousness is.

The illusionist position is that phenomenal properties are cognitive illusions generated by the brain. If a zombie with developed cognitive abilities can mistakenlythink it has phenomenal consciousness, how do we know that this is not the case with ourselves, as well? For the illusionist, this is exactly the predicament we are in, albeit we are zombies with a rich inner life (Frankish, 2017)—whatever that means. The illusionist takes the purported scientific intractability of phenomenal consciousness to be evidence against phenomenal consciousness. "Illusionists deny that experiences have phenomenal properties and focus on explaining why they seem to have them" (Frankish, 2017, p. 18). While this position might seem to be counterintuitive, it saves a conservative understanding of physics and obviates any call for exotic properties of the universe, as Chalmers (2017) argues for. As Graziano et al. (2020) state:

> "'I know I have an experience because, Dude, I'm experiencing it right now.' Every argument in favour of the literal reality of subjective experience . . . boils down sooner or later to that logic. But the logic is circular. It is literally, 'X is true because X is true.' If that is not a machine stuck in a logic loop, we don't know what is" (Graziano et al., 2020, p. 8).

The problem, however, is that it's not clear what the claim that phenomenal consciousness is an illusion means. What exactly does "illusion" mean in this context? To be clear, the illusionist is not denying that you see, e.g., a red rose in front of you. You do "see" it, but it only *seems* like it has phenomenal properties. You have non-phenomenal access to the perceptual representation, sufficient to enable a phenomenal judgment about that representation (e.g., "I see a red rose"). This contention runs in the face of all our intuitions, but the illusionist claims those intuitions are illusory. Frankish (2017) states that these *seeming*-properties are quasi-phenomenal properties, which are physical properties that give the illusion of phenomenal properties. But are quasi-phenomenal properties any less mysterious than phenomenal properties? How is it that seeming to be phenomenally conscious is not just being phenomenally conscious?

One way to address these problems and understand the illusionist position is to understand them as agreeing with Chalmers about phenomenal consciousness's privacy, and that zombies are logically possible. Then, the paradox of phenomenal judgment already notes that just such a functionally conscious cognitive system could produce phenomenal judgments without having phenomenal consciousness. However, illusionists take this to be evidence against phenomenal consciousness. A purely physical system creates phenomenal judgments, therefore there are nothing but purely physical processes involved (and hence, no qualia). However, this position is problematic. Only if phenomenal properties are private could there be such a paradoxical situation of having phenomenal judgments about phenomenal consciousness, yet without any phenomenal consciousness. If we can argue *against* the privacy of phenomenal properties, then we can escape the trap into which both the dualist and illusionist fall.

We interpret the dualist and illusionist extremes as unfortunate consequences of a mistaken view of naturalism. The illusionist's commitment to naturalism leads them to exclude supposedly non-natural properties like qualia. The scientific dualist is also committed to naturalism, but takes it that the current inventory of nature is simply incomplete and that there must be a new, exotic fundamental property in the universe, that of phenomenality.

The Relativistic Approach

A common thread connecting both extremes of dualism and illusionism is that both assume that phenomenal consciousness is an absolute phenomenon, wherein an object O evinces either property P or $\neg P$. We will show that we need to abandon this assumption. The relativistic principle in modern physics posits a universe in which for many properties an object O evinces either property P or $\neg P$ *with respect to some observer X*. In such a situation, there is no one answer to the question of whether object O has property P or not. We propose a novel relativistic theory of consciousness in which consciousness is not an absolute property but a relative one. This approach eschews both extremes of illusionism and dualism. The relativistic theory of consciousness will show that phenomenal consciousness is neither an illusion created by a "machine stuck in a logic loop" nor a unique fundamental property of the universe. It will give a coherent answer to the question of the (supposed) privacy of phenomenal consciousness, will bridge the explanatory gap, and will provide a solution to the hard problem based in relativistic physics. General notions of this approach can be found in some dual aspect monisms such as Max Velmans' reflexive monism. According to Velmans (2009, p. 298), "[i]ndividual conscious representations are perspectival." Here, however, we develop a physical theory of consciousness as a relativistic phenomenon and formalize the perspectival relations in light of the relativistic principle. To do that, in the following section, we will develop more formally the *relativistic principle* and introduce the *equivalence principle of consciousness*.

In physics, relativity means that different observers from different frames of reference will nevertheless measure the same laws of nature. If, for example, one observer is in a closed room in a building and the other observer is in a closed room in a ship (one moving smoothly enough on calm water), then the observer in the ship would not be able to tell whether the ship is moving or stationary. Each will obtain the same results for any experiment that tries to determine whether they are moving or not. For both of them, the laws of nature will be the same, and each will conclude that they are stationary. For example, if they throw a ball toward the room's ceiling, each will determine that the ball will return directly into their hands (because the ship moves with constant velocity and because of Newton's First Law, the ball will preserve the velocity of the ship while in the air, and will propagate forward with the same pace as the ship. As a result, it will fall directly into the observer's hands). There will be no difference in the results of each observer's measurements, trying to establish whether they are stationary or not. They will conclude that they have the same laws of nature currently in force, causing the same results. These results will be the results of a stationaryobserver and thus both of them will conclude that they are at rest. Because each of them will conclude that they are the stationary one, they will not agree about one another's status. Each of them will conclude that the other is the one that moves (common sense will tell the observer in the ship that they are the one moving, but imagine an observer locked on the ship in a room with no windows. Such an observer cannot observe the outside world. This kind of observers will conclude that they are stationary because velocity is relativistic).

To state that consciousness is a relativistic phenomenon is to state that there are observers in different *cognitive frames of reference,* yet they will nevertheless measure the same laws of nature currently in force and the same phenomenon of consciousness in their different frames. We will start with an equivalence principle between a conscious agent, like a human being, and a zombie agent, like an advanced artificial cognitive system. As a result of this equivalence, we will show that if the relativistic principle is true, then zombies are not possible. Instead, every purported zombie will actually have phenomenal consciousness and any system with adequate functional consciousness will exhibit phenomenal consciousness from the first-person cognitive frame of reference. Others have similarly claimed that zombies are physically impossible (Brown, 2010; Dennett, 1995; Frankish, 2007; Nagel, 2012), but our aim here is to show why that is according to the relativistic principle. As a result of this equivalence, observations of consciousness fundamentally depend on the observer's cognitive frame of reference. The first-person cognitive frame of reference is the perspective of the cognitive system itself (Solms, 2021). The third-person cognitive frame of reference is the perspective of any external observer of that cognitive system. Phenomenal consciousness is only seemingly private because in order to measure it one needs to be in the appropriate cognitive frame of reference. It is not a simple transformation to change from a third-person cognitive frame of reference to the first-person frame, but in principle it can be done, and hence phenomenal consciousness isn't private anymore. We avoid the term "first-person perspective" because of its occasional association with immaterial views of consciousness; cognitive frames of reference refer to physical systems capable of representing and manipulating inputs. These systems have physical positions in space and time and instantiate distinct dynamics.

In section "The Equivalence Principle of Consciousness: Mathematical Description," we will show that from its own first-person cognitive frame of reference, the observer will observe phenomenal consciousness, but any other observer in a third-person cognitive frame of reference will observe only the physical substrates that underlie qualia and eidetic structures. The illusionist mistake is to argue that the third-person cognitive frame of reference is *the* proper perspective. To be clear, the first-person cognitive frame of reference is still a physical location in space and time (it is not immaterial); it is just the position and the dynamics of the cognitive system itself. As we will see, this is the position from which phenomenal consciousness can be observed. The principle of relativity tells us that there is no privileged perspective in the universe. Rather, we will get different measurements depending on the observer's position. Since consciousness is relativistic, we get different measurements depending on whether the observer occupies or is external to the cognitive system in question. Both the first-person and third-person cognitive frames of reference describe the same reality from two different points of view, and we cannot prefer one point of view upon the other (Solms, 2021). The dualist mistake is to argue that phenomenal consciousness is private. For any relativistic phenomenon there is a formal transformation between the observers of different frames of reference, meaning that both frames can be accessible to every observer with the right transformations. Consciousness as a relativistic phenomenon also has such transformation rules. We will describe the transformations between first-person (i.e., phenomenological) and third-person (i.e., laboratory point of view) cognitive frames of reference. There are several consequences of these transformations. First, qualia and eidetic structures are not private. Rather, they only *appear* private, because in order to measure them one needs to be in the appropriate cognitive frame of reference, i.e., within the perspective of the cognitive system in question. We can use these transformations to answer questions like, "What is it like to be someone else?" Because of the transformations, results that we obtain from third-person methodology should be isomorphic to first-person structures. Isomorphism between two elements means that they have the same mathematical form and there is a transformation between them that preserves this form. Equality is when two objects are exactly the same, and everything that is true about one object is also true about the other. However, an isomorphism implies that everything that is true about *some* properties of one object's structure is also true about the other. In section "The Equivalence Principle of Consciousness: Mathematical Description," we will show that this is the case with measurements obtained from first-person and third-person frames of reference. We will show that this isomorphism is a direct result of the relativistic principle and the notion that phenomenal judgements and phenomenal structures are two sides of the same underlying reality. All that separates them are different kinds of measurements (causing different kinds of properties). An unintuitive consequence of the relativistic theory is that the opposite is also true, and first-person structures also bear formal equivalence to third-person structures. We advocate for interdisciplinary work between philosophers and cognitive neuroscientists in exploring this consequence.

THE EQUIVALENCE PRINCIPLE OF CONSCIOUSNESS

The Principle of Relativity and the Equivalence Principle

Our task is to establish the equivalence principle of consciousness, namely, that qualitative and quantitative aspects of consciousness are formally equivalent. We start by establishing the equivalence between conscious humans and zombies, and then we expand that equivalence to all structures of functional consciousness. We begin with a philosophical defense of the equivalence principle and then develop a mathematical formalization. We must first present the principle of relativity and the equivalence principle in physics. Later, we will use these examples to develop a new equivalence principle and a new transformation for consciousness. To be more formal than earlier, the principle of relativity is the requirement that the equations describing the laws of physics have the same form in all admissible frames of reference (Møller, 1952). In physics, several relativistic phenomena are well-known, such as velocity and time, and the equivalence principle between uniformly accelerated system and a system under a uniform gravitational field. Let's examine two examples with the help of two observers, Alice and Bob.

In the first example of a simple Galilean transformation, Alice is standing on a train platform and measures the velocity of

Bob, who is standing inside a moving train. Meanwhile, Bob simultaneously measures his own velocity. As the train moves with constant velocity, we know that the laws of nature are the same for both Alice's and Bob's frames of reference (and thus that the equations describing the laws of physics have the same form in both frames of reference; Einstein et al., 1923/1952, p. 111). According to his measurements, Bob will conclude that he'sstationary, and that Alice and her platform are moving. However, Alice will respond that Bob is mistaken, and that *she* is stationary while Bob and the train are moving. At some point, Alice might say to Bob that he has an illusion that he is stationary and that she is moving. After all, it is not commonsensical that she, along with the platform and the whole world, are moving. Still, although it doesn't seem to make common sense, in terms of physics all of Bob's measurements will be consistent with him being stationary and Alice being the one who moves. In a relativistic universe, we cannot determine who is moving and who is stationary because all experiment results are the same whether the system is moving with constant velocity or at rest. Befuddled, Bob might create an elaborate argument to the effect that his measurement of being stationary is a private measurement that Alice just can't observe. But, of course, both of their measurements are correct. The answer depends on the frame of reference of the observer. Yet, both of them draw mistaken conclusions from their correct measurements. Bob has no illusions, and his measurements are not private. Velocity is simply a relativistic phenomenon and there is no answer to the question of what the velocity of any given body is *without reference to some observer*. Their mistakes are derived from the incorrect assumption that velocity is an absolute phenomenon. As counterintuitive as it may seem, the relativity principle tells us that Bob is not the one who is "really" moving. Alice's perspective is not some absolute, "correct" perspective that sets the standard for measurement, although we may think that way in our commonsense folk physics (Forshaw and Smith, 2014). Alice is moving relative to Bob's perspective, and Bob is moving relative to Alice's perspective. Furthermore, the fact that Alice and Bob agree about all the results from their own measurements means that these two frames of reference are physically equivalent (i.e., they have the same laws of nature in force and cannot be distinguished by any experiment). Because of this physical equivalence, they will not agree about who is at rest and who is moving.

Later, Einstein extended the relativity principle by creating special relativity theory. The Galilean transformation showed that there is a transformation between all frames of reference that have constant velocity relative to each other (inertial frames). Einstein extended this transformation and created the Lorentz transformation, which describes more accurately the relativistic principle. enabling us to move from measurements in one inertial frame of reference to measurements in another inertial frame of reference, even if their velocity is near the speed of light (Einstein, 1905; Forshaw and Smith, 2014). According to the transformation, each observer can change frames of reference to any other inertial frame by changing the velocity of the system. The transformation equation ensures that in each frame we'll get the correct values that the system will measure. For example, the observer will always measure that its own system is at rest and that it is the origin of the axes. Indeed, this is the outcome of the transformation equation for the observer's own system. Mathematically, a frame of reference consists of an abstract coordinate system, and (in tensor formulation) we denote each frame of reference by a different Greek letter (and by adding a prime symbol), usually by μ and ν, which are indexes for the elements of the vectors in each coordinate system. The Lorentz transformation is denoted by $\Lambda^{\mu\nu}$. It is a matrix that gets elements of a vector in one frame of reference (x'^{μ}) and gives back the elements of a vector in another frame of reference (x^{ν}):

$$x^{\nu} = \Lambda^{\nu}_{\mu} x'^{\mu} \tag{1}$$

This equation coheres with the relativistic principle. That is, it describes the laws of physics with the same form in all admissible frames of reference. This form will stay the same regardless of which frame we choose. It allows us to switch frames of reference and get the measured result in one frame, depending only on the other frame and the relative velocity between the two.

The second example comes from Albert Einstein's (1907) observation of the equivalence principle between a uniformly accelerated system (like a rocket) and a system under a uniform gravitational field (like the Earth). Here, Einstein extended the relativity principle even more, not only to constant velocity but also to acceleration and gravity (for local measurements that measure the laws of nature near the observer). He started with two different frames of reference that have the exact same results from all measurements made in their frames of reference. Then he used the relativity principle, concluding that because they cannot be distinguished by any local experiment they are equivalent and have the same laws of nature. Lastly, he concluded that because of the equivalence, we can infer that phenomena happening in one frame of reference will also happen in the other. Now, assume that Alice is skydiving and is in freefall in the Earth's atmosphere, while Bob is floating in outer space inside his spaceship. Although Alice is falling and Bob is stationary in space, both will still obtain the same results of every local experiment they might do. For example, and if they were to release a ball from their hand, both of them will measure the ball floating beside of them (in freefall, all bodies fall with the same acceleration and appear to be stationary relative to one another—this is why movies sometimes use airplanes in freefall to simulate outer space). Although they are in different physical scenarios, both will infer that they are floating at rest. Einstein concluded that because they would measure the same results, there is an equivalence between the two systems. In an equivalence state, we cannot distinguish between the systems by any measurement, and thus a system under either gravity or acceleration will have the same laws of nature in force, described by the same equations regardless of which frame it is.

Because of this equivalence, we can infer physical laws from one system to the other. For example, from knowing about the redshift effect of light in accelerated systems with no gravitational force, Einstein predicted that there should be also a redshift effect of light in the presence of a gravitational field as if it were an accelerated system. This phenomenon was later confirmed (Pound and Rebka, 1960). According to Einstein (1911),

> "By assuming this [the equivalence principle] to be so, we arrive at a principle which. . .has great heuristic importance. For by theoretical consideration of processes which take place relative to a system of reference with uniform acceleration, we obtain information as to the behavior of processes in a homogeneous gravitational field" (p. 899).

Next, we argue that phenomenal consciousness is a relativistic physical phenomenon just like velocity. This allows us to dissolve the hard problem by letting consciousness be relativistic instead of absolute. Moreover, we develop a similar transformation for phenomenal properties between first-person and third-person perspectives. The transformation describes phenomenal consciousness with the same form in all admissible cognitive frames of reference and thus satisfying the relativistic principle.

The Equivalence Principle of Consciousness: Conceptual Argument

Before we begin the argument, it is essential to elaborate on what 'observer' and 'measurement' mean when applying relativistic physics to cognitive science. In relativity theory, an observer is a frame of reference from which a set of physical objects or events are being measured locally. In our case, let's define a '*cognitive frame of reference*'. This is the perspective of a specific cognitive system from which a set of physical objects and events are being measured. Cognitive frame of reference is being determined by the dynamics of the cognitive system (for more details, see section "The Equivalence Principle of Consciousness: Mathematical Description").

We use the term 'measurement' in as general a way as possible, from a physical point of view, such that a measurement can occur between two particles like an electron and a proton. Measurement is an interaction that causes a result in the world. The result is the measured property, and this measured property is new information in the system. For example, when a cognitive system measures an apple, it means that there is a physical interaction between the cognitive system and the apple. As a result, the cognitive system will recognize that this is an apple (e.g., the interaction may be via light and the result of it will be activation of retinal cells which eventually, after sufficient interactions, will lead to the recognition of the apple). In the case of cognitive systems, there are measurements of mental states. This kind of measurement means that the cognitive system interacts with a content-bearing cognitive state, like a representation, using interactions between different parts of the system. It is a strictly physical and public process (accessible for everybody with the right tools). As a result of this definition, the starting point of the argument is with measurements that are non-controversial, i.e., measurements that are public. These measurements include two types, measurements of behavioral reports and measurements of neural representations (like the phenomenal judgement-representations). For example, physical interaction of light between an apple and a cognitive system causes, in the end of a long process of interactions, the activation of phenomenal judgement-representation of an apple and possibly even a behavioral report ("I see an apple").

Let us start from the naturalistic assumption that phenomenal consciousness should have some kind of physical explanation. The physicalism we assume includes matter, energy, forces, fields, space, time, and so forth, and might include new elements still undiscovered by physics. Panpsychism, naturalistic dualism, and illusionism all fall under such a broad physicalism. (It might be, for example, that physics hasn't discovered yet that there is a basic private phenomenal element in addition to the observed known elements. If this element exists, it needs to be part of the broad physicalism.) In addition, let's assume that the principle of relativity holds, i.e., all physical laws in force should be the same in different frames of reference, provided these frames of references agree about all the results of their measurements. Since we have accepted that consciousness is physical (in the broad sense), we can obtain a new equivalence principle for consciousness. Let's assume two agents—Alice, a conscious human being—and a zombie in the form of a complex, artificial cognitive system that delusionally claims to have phenomenal consciousness. Let's call this artificially intelligent zombie "Artificial Learning Intelligent Conscious Entity," or ALICE. ALICE is a very sophisticated AI. It has the capacity to receive inputs from the environment, learn, represent, store and retrieve representations, focus on relevant information, and integrate information in such a manner that it can use representations to achieve human-like cognitive capabilities.

ALICE has direct access to perceptual information and to some of its own cognitive contents, a developed concept of self, the capacity to reflect by creating representations of its internal processes and higher-order representations, and can create outputs and behavioral responses (it has a language system and the ability to communicate). In fact, it was created to emulate Alice. It has the same representations, memories, and dynamics as Alice's cognitive system and as a result, it has the same behavioral responses as Alice. But ALICE is a zombie (we assume) and doesn't have phenomenal consciousness. On the other hand, conscious Alice will agree with all of zombie ALICE's phenomenal judgments. After sufficient time to practice, ALICE will be able to produce phenomenal judgments and reports nearly identical to those of Alice. After enough representational manipulation, ALICE can say, e.g., "I see a fresh little madeleine, it looks good and now I want to eat it because it makes me happy." It can also reflect about the experience it just had and might say something like, "I just had a tasty madeleine cake. It reminded me of my childhood, like Proust. There's nothing more I can add to describe the taste, it's ineffable."

Alice and ALICE will agree about all measurements and observations they can perform, whether it's a measurement of their behavior and verbalizations, or even an "inner" observation about their own judgments of their experience, thoughts and feelings. They will not find any measurement that differs between them, although Alice is conscious, and ALICE is not. For example, they can use a Boolean operation (with yes/no output) to compare their phenomenal judgments (e.g., do both agree that they see a madeleine and that it's tasty?). Now, let's follow in Einstein's footsteps concerning the equivalence principle between a uniformly accelerated system and a system under a uniform gravitational field. Alice and ALICE are two different observers, and the fact that they obtain the same measurements agrees with the conditions of the relativity principle. Accordingly, because both of them completely agree upon all measurements, they are governed by the same physical laws. More precisely, these two observers are cognitive systems, each with its own cognitive frame of reference. For the same input, both frames of reference agree about the outcome of all their measurements, and thus both currently have the same physical laws in force. As a result, these two systems are equivalent to each other in all physical aspects and we can infer physical laws from one system to the other. According to the naturalistic assumption, phenomenal consciousness is part of physics, so the equivalence between the systems applies also to phenomenal structures. Consequently, we can infer that if Alice has phenomenal properties, then ALICE also must have them. Both Alice and ALICE must have phenomenal consciousness! ALICE cannot be a zombie, like we initially assumed, because their systems are physically equivalent. This equivalence makes it impossible for us to speak of the existence of absolute phenomenal properties in the human frame of reference, just as the theory of relativity forbids us to talk of the absolute velocity in a system. For, by knowing that there is phenomenal consciousness in the human frame of reference and by using the broad physicalist premise, we conclude that there are physical laws that enable phenomenal consciousness in the human frame of reference. Because of the equivalence principle, we can infer that the same physical laws will be present also in the supposed "zombie" cognitive system's frame of reference. The conclusion is that if there is phenomenal consciousness in the human frame of reference, then the "zombie" cognitive system's frame of reference must also harbor phenomenal consciousness.

We started from the premise that ALICE is a zombie and concluded that it must have phenomenal consciousness. One of our premises is wrong: either the broad physicalism, the relativistic principle, or the existence of zombies. Most likely the latter is the odd man out, because we can explain these supposed "zombies" using a relativistic, physicalist framework. As a result, although we started from a very broad notion of physicalism and an assumption that the human has phenomenal consciousness and the "zombie" cognitive system does not, the relativity principle forces us to treat phenomenal structures as relativistic. According to the relativity principle, there is no absolute frame of reference; there are only different observers that obtain different measurements. If the observers obtain the same measurements, there cannot be anything else that influences them. There is nothing over and above the observers (no God's-eye-view), and if they observe that they have phenomenal properties, then they have phenomenal consciousness. Because there is nothing over and above the observer, we can generalize this result even further for every cognitive system that has phenomenal judgments: Any two arbitrary cognitive observers that create phenomenal judgment-representations also have phenomenal consciousness. Zombies cannot exist (assuming their cognitive systems create phenomenal judgments). In other words, we obtain an equivalence between functional consciousness (which creates phenomenal judgments) and phenomenal consciousness. Notice that even if we start from the naturalistic dualism of Chalmers and assume a broader physics including phenomenal elements alongside other aspects in the universe, the relativistic principle still forces this kind of broad physics to have the same consequences, viz., that zombies cannot exist and that there is an equivalence between phenomenal judgement-representations and phenomenal properties. (Formally, phenomenal judgement-representations and phenomenal properties are isomorphic. They have the same mathematical form and there is a transformation between them that preserves this form. See section "The Equivalence Principle of Consciousness: Mathematical Description"). As a result, the relativistic principle undermines the dualist approach altogether.

In the next section, we develop the mathematical proof of the argument. The mathematical description elaborates the fine details of the theory and reveals new insights. Notice that a mathematical background is not necessary to understand this section, as every step includes a comprehensive explanation.

The Equivalence Principle of Consciousness: Mathematical Description

In this section we develop a mathematical description of consciousness as a relativistic phenomenon. In the beginning of the argument, we apply the relativistic principle only to publicly-observable measurements. Then, we show that we can also apply the relativistic principle to phenomenal properties and structures. To that end, we use two different arguments. Then, using the relativistic principle, we prove the equivalence principle between a conscious human agent and a purported "zombie." Then, we expand that equivalence to all cognitive frames of reference having functional consciousness (i.e., the cause of phenomenal judgments). As a result, we prove that phenomenal judgements are equal to phenomenal properties in the cognitive frame of reference that generates them. Then, we describe the difference between the first-person and third-person perspectives and develop a transformation between them and between measurements of any cognitive frames of reference that have phenomenal consciousness. We show that this transformation preserves the form of the equation regardless of which frame we choose, and thus satisfies the relativity principle. That is, it describes the laws of physics with the same form in all admissible frames of reference.

The Three-Tier Information Processing Model for Cognitive Systems

Let's once more assume two agents, Alice, a conscious human agent, and ALICE, a "zombie" artificial cognitive system (at least, supposedly) with phenomenal judgments. These systems parallel Chalmers' thought experiment: ALICE is the cognitive duplicate of Alice, but lacks (we suppose) phenomenal consciousness. In order to develop a fully mathematical description, we need to know the exact equations of the cognitive systems in question.

Unfortunately, to date there is no complete mathematical description of a human cognitive system, let alone an artificial one that can mimic human cognition. Instead, we use a general, simplified version, a mathematical toy model, to describe a cognitive system that creates phenomenal judgments. We use a three-tier information processing model that divides a cognitive system (*S*) into three parts: sensation (*T*), perception (*P*), and cognition (Dretske, 1978, 2003; Pageler, 2011). Sensation picks up information from the world and transforms it, *via* transduction, into neural signals for the brain. Later, perceptual states are constructed *via* coding and representations until a percept is created. The point at which cognitive processing is thought to occur is when a percept is made available to certain operations, such as recognition, recall, learning, or rational inference (Kanizsa, 1985; Pylyshyn, 2003). Cognition has many different operations, but our main interest is the module that specializes in functional consciousness (and that creates phenomenal judgments), which we label *C*. After recognition, this functional consciousness module needs to integrate the input and emotional information, the state of the system, the reaction of the system to the input, and self-related information in order to create representations of phenomenal judgments. We can summarize this model with the following equation:

$$S = (T, P, C) \tag{2}$$

where *T* is the sensation subsystem, *P* is the perception subsystem, *C* is the functional consciousness subsystem of cognition, which is ultimately responsible for phenomenal judgments (we're not considering other cognitive subsystems here). Finally, *S* is the entire cognitive system as a whole. These subsystems are a 3-tuple that form the cognitive system in a sequential manner. From *T* to *P* and to *C*. The three-tier model yields:

$$\vec{t^{i}_{x_t}} = T\left(\vec{x}_t\right) \tag{3}$$

$$\vec{p}_{x_t} = P\left(\vec{t^{i}_{x_t}}\right) \tag{4}$$

$$\vec{c}_{x_t} = C\left(\vec{p}_{x_t}\right) \rightarrow \vec{c}_{x_t} = S(\vec{x}_t) \tag{5}$$

where $\vec{x}_t$ is an input from the physical space outside the cognitive system that interacts with the cognitive system at time *t*. The input has physical properties, y_j (E.g., temperature, position, velocity, etc. j = 1,2,3... is an index of the properties). These physical properties can be detected by the sensors of the sensation subsystem (e.g., the input can be an apple and the cognitive system interacts with it by a beam of light that reached its retina at time *t*. Using photoreceptors in the retina, the sensation subsystem can detect physical properties of the apple like its color.) *T* is the sensation subsystem, and $\vec{t^{i}_{x_t}}$ are the outputs of the subsystem, i = 1,2,3...n is an index for each sensory module (i.e., $\vec{t^{1}_{x_t}}$ is the output vector of the sensation subsystem for visual information, $\vec{t^{2}_{x_t}}$ for auditory information, $\vec{t^{3}_{x_t}}$ for interoception information etc.). Using the sensation subsystem, the cognitive system picks up information from the world and transforms it into neural patterns. Each pattern is a different possible state in the state space of the cognitive system. In this state space every degree of freedom (variable) of the dynamic of the system is represented as an axis of a multidimensional space and each point is a different state (e.g., the axes can represent firing rate, membrane potential, position, etc.). After a learning period a subset of these states is formed and being used by the sensation subsystem to encode inputs. This is the state space of the sensation subsystem, and it is part of the state space of the whole cognitive system. $\vec{t^{i}_{x_t}}$ are vectors that point to such states in the state space of the sensation subsystem. There is a correspondence between the states $\vec{t^{i}_{x_t}}$ and the captured physical properties y_j (e.g., correspondence between the color of an apple and the neural pattern that it causes in the system. light from a red apple activates photoreceptors which create a neural pattern. In this pattern, cones sensitive to red frequencies will fire more.) *P* is the perception subsystem. It gets $\vec{t^{i}_{x_t}}$ as inputs from the sensation subsystem and returns output $\vec{p}_{x_t}$, the percept of the input $\vec{x}_t$. The perception subsystem creates representations. These are gists of important information about the input $\vec{x}_t$ that the cognitive system can use even when the input is absent. the percept $\vec{p}_{x_t}$ is a representation unifying information about the current input from all sensory modules. To that end, the perception subsystem using yet another subset of states from the state space. This is the state space of the perception subsystem, and it is part of the state space of the whole cognitive system. $\vec{p}_{x_t}$ is a vector that points to such a state in the state space of the perception subsystem. There is a correspondence between the states and the operations that can act on these states and between the captured physical properties y_j (e.g., one operation on states can be addition of two states to get a third state in the state space of the perception subsystem. In our toy model, this operation can correspond to integration of different physical properties. A red round shaped object, for example, can be a state that is a summation of the red and round states in the state space of the perception subsystem.) *C* is the functional consciousness subsystem that creates phenomenal judgments (representations that are not phenomenal states themselves, but they are about phenomenology. see introduction for details). It gets percepts as inputs from the perception subsystem and returns phenomenal judgment representation, $\vec{c}_{x_t}$ as output. As before, the functional consciousness subsystem using yet another subset of states from the state space. $\vec{c}_{x_t}$ is a vector that points to such a state in the state space of the functional consciousness subsystem. There is a correspondence between the operations that can act on these states in the functional consciousness subsystem state space and between the captured physical properties y_j. Combining these equations with eq. 2, the conclusion is that the cognitive system *S*, which uses the sensation, perception and functional consciousness subsystem sequentially, has a state space that is a combination of the three subsystems state spaces. It gets as input, $\vec{x}_t$ from the physical space, and returns as output, $\vec{c}_{x_t}$, the phenomenal judgment representation, from the state space of the cognitive system. To sum, when Alice sees a red apple, for example, her perception subsystem creates a representation of sensory information, a unification of representations for "red," "round," and possibly a smell and texture. This integrated representation is her current percept. Later, the functional consciousness subsystem will recognize the object as an "apple" and will create an integrated representation of all relevant information about this red apple. This phenomenal judgment can be, for example, "good-looking, red, and roundish apple," which can later be reported. Such a cognitive system S needs to integrate information from previous layers and from different modules. It needs an attention module (W) that can change the weights of representations and filter information to focus on relevant information. It needs a long-term memory module (M) to store and retrieve representations. It also needs an emotion evaluation module (E), which assesses information and creates preferences, rewards, and avoidances (e.g., does the input have a positive or negative affective valence?). Other modules could include an affordance module (A) that detects possibilities for action in the environment according to the input, and a self-module (I) that collects self-related information to create a unified representation of the self. Lastly, the cognitive system needs modules to create outputs and behavioral responses. In particular, it has a language module (L) that creates lexical, syntactic, and semantic representations and responses that the functional consciousness subsystem (C) can use. Adding them to Equations 4, 5, we obtain:

$$\vec{p}_{x_t} = P\left(\vec{t^{i}_{x_t}}, E, W, M\right) \quad (6)$$

the perception subsystem P receives as variables the outputs of the sensation subsystem $\vec{t^{i}_{x_t}}$ and uses the long-term memory module M, the attention module W, and the evaluation module E to create a percept $\vec{p}_{x_t}$ of the input $\vec{x}_t$, a unified representation of perception from all sensory modules. This yields:

$$\vec{c}_{x_t} = C\left(\vec{p}_{x_t}, E, W, M, A, I\right) \quad (7)$$

the functional consciousness subsystem C receives as a variable the output of the perception subsystem $\vec{p}_{x_t}$ and uses the long-term memory module M, the attention module W, the evaluation module E, the affordance module A and the self-module I to create a phenomenal judgment $\vec{c}_{x_t}$ (a complex representation that carries content about phenomenology) concerning the input $\vec{x}_t$. We can also add the language module for the cognitive system, so that S can understand and answer questions:

$$\vec{p}_l = L\left(\vec{a}_l, E, W, M\right) \quad (8)$$

where L is the language module, $\vec{a}_l$ is the input of the language module, a low-level processed representation from the perception subsystem (without loss of generality, let us assume it to be auditory information. It can also be a low-level representation of other sensory modules). The language module L gets a low-level representation from the perception subsystem $\vec{a}_l$ and uses the long-term memory module M, the attention module W, and the evaluation module E to create several representations. Semantic representations for words, syntactic representations, and eventually sentence comprehension representation, $\vec{p}_l$. As before, the language module using yet another subset of states from the state space. $\vec{p}_l$ is a vector that points to such a state in the state space of the language module. The functional consciousness subsystem C uses the output of the language module to perform its tasks and create phenomenal judgments (answering a question about a red apple, for example):

$$\vec{c}_{p_{l,p_{x_t}}} = C\left(\vec{p}_l, \vec{p}_{x_t}, E, W, M, A, I\right) \quad (9)$$

Now we add $\vec{p}_l$ to Equation 7, which is the sentence comprehension representation (a question, for example) as another variable of C. The answer, $\vec{c}_{p_{l,p_{x_t}}}$, is sent to the language module, this time creating a linguistic response:

$$\vec{l}_{p_{l,x_t}} = L\left(\vec{c}_{p_{l,p_{x_t}}}, E, W, M\right) \quad (10)$$

where L is the language module and $\vec{l}_{p_{l,x_t}}$ is a vector of the linguistic response. As before, the language module using yet another subset of states from the state space. $\vec{l}_{p_{l,x_t}}$ is a vector that points to such a state in the state space of the language module. This state is a complex neural pattern caused by motor neurons that innervate muscle fibers. The pattern causes contraction patterns in the muscles to create a linguistic response. The response is according to $\vec{c}_{p_{l,p_{x_t}}}$, the phenomenal judgement representation that captures both the question and the answer of C. For example, suppose that Alice asks ALICE what she sees, and ALICE sees a red apple. The visual information will be transduced and processed in the perception subsystem until a percept is created (equation 6). The language system creates a representation of the question, $\vec{p}_l$ = "what do you see?" (eq. 8) and the functional consciousness subsystem uses the representation of the question, $\vec{p}_l$ and the percept of the apple, $\vec{p}_{x_t}$, as variables to build a phenomenal judgement to answer the question, $\vec{c}_{p_{l,p_{x_t}}}$ (eq. 9). Finally, the language module uses this representation to produce a proper linguistic response, $\vec{l}_{p_{l,x_t}}$ = "I see an apple" (eq. 10).

The Equivalence Principle of Consciousness

In parallel with Chalmers' (1996) assumptions about zombies, Alice and ALICE have equivalent cognitive systems. But Alice, being human, also has phenomenal consciousness (without loss of generality, let's assume Alice has a quale Q_{x_t} about input $\vec{x}_t$. We will not assume any inner structure for the quale). After we establish the equations for the cognitive system, we can use them to prove the equivalence principle for consciousness. Let's start with general considerations about cognitive frames of reference. A cognitive frame of reference is determined by the dynamics of a cognitive system (represented by Equations 1–10). Because the equations create and use representations, the representations are an inseparable part of the dynamics of a cognitive system. If two cognitive systems have different dynamics and representations, they are in different cognitive frames of reference. But if they have the same dynamics and representations, then they are in the same cognitive frame of reference. We will denote cognitive frames of reference by μ, and ν.

We started with the assumption that ALICE is a zombie. Hence, ALICE will obtain the same behavioral reports (i.e., $\vec{l}_{p_{l,x_t}}$) and the same phenomenal judgments ($\vec{c}_{x_t}$) as Alice, who also has phenomenal consciousness (note that for this claim to be true, we idealize the situation as if Alice and ALICE were standing in the same spatial position. Hence, they will obtain the same phenomenal judgments. We can neglect minor differences arising from their different spatial positions, as long as they are standing near each other). In fact, Alice and ALICE have the same exact cognitive structure and dynamics (i.e., the same subsystems, modules, dynamics, and representations). As a result, they are in the same cognitive frame of reference and for each question they will have the same responses and the same phenomenal judgments. We can formalize it as such:

$$<Alice>^{\nu} = \widetilde{\Lambda}^{\nu\mu} <ALICE>^{\mu} \quad (11)$$

where ν is Alice's cognitive frame of reference, μ is ALICE's cognitive frame of reference, and $< \;\; >$ is a symbol for the set of all results of measurements of a system (i.e., $<$ Alice $>$ means the set of results of measurements conducted by Alice). $\widetilde{\Lambda}^{\nu\mu}$ is a general transformation from the set of results of measurements of frame μ to the corresponding set of results of measurements of frame ν. Alice and ALICE can measure or observe both their behavioral reports and their phenomenal judgments, and then compare the results. They will discover that their results are the same. For the same input, they will not find any measurement about themselves to distinguish between them (again, the fact that they stand in slightly different spatial positions can be neglected because any reasonable cognitive system takes different positions into account in order to build percepts correctly, regardless of the system's position. For example, the cognitive system needs to recognize an apple regardless of the angle between the apple and the system. Objects should be invariant structures across changing conditions.) According to the relativistic principle, because they obtain the same results about themselves, they have the same physical laws currently in force in their respective frames of reference. There is an equivalence between the two systems, and we can infer physical laws and phenomena from one system to the other. For example, we know that

$$\cup \vec{l}^{\nu}_{p_{l,x_t}} = \widetilde{\Lambda}^{\nu\mu} \cup \vec{l}^{\mu}_{p_{l,x_t}} \quad (12)$$

There is a transformation from the unification of all linguistic responses of Alice, $\cup \vec{l}^{\nu}_{p_{l,x_t}}$, to the unification of all linguistic responses of ALICE, $\cup \vec{l}^{\mu}_{p_{l,x_t}}$ (provided that both were asked the same questions, $\vec{p}_l$). Let's consider a specific linguistic response. From Equation 11 we obtain:

$$\vec{l}^{\nu}_{p_{l,x_t}} = \widetilde{\Lambda}^{\nu\mu} m^{\mu}, \quad (13)$$

where m^{μ} is a specific measurement result of ALICE, $m^{\mu} \in <ALICE>$. $\vec{l}^{\nu}_{p_{l,x_t}}$ is a specific linguistic response of Alice. There is a transformation from the linguistic response of Alice that she will measure (i.e., that she will observe about herself) in her cognitive frame of reference to a specific measurement that ALICE will obtain about herself in her frame of reference. Now we can ask, what is the specific measurement in ALICE's frame of reference? Because of the relativistic principle we know that the frames of reference of Alice and ALICE are equivalent, and hence we just need to replace ν with μ:

$$\vec{l}^{\nu}_{p_{l,x_t}} = \delta^{\nu\mu} m^{\mu} \quad (14)$$

$$\vec{l}^{\mu}_{p_{l,x_t}} = m^{\mu} \quad (15)$$

We use a special function known as the delta function, $\delta^{\nu\mu}$, to transform from frame ν to frame μ. This function equals 0 if $\nu \neq \mu$, and equals 1 if $\nu = \mu$. In other words, the only solution that is different from 0 for Equation 14 is when $\nu = \mu$. Now we can plug in 1 instead of $\delta^{\nu\mu}$ and substitute μ for ν on the left-hand side of Equation 15, with the result that ALICE will measure $\vec{l}^{\mu}_{p_{l,x_t}}$. Now we can plug the result of Equation 15 back in Equation 13, which gives us a similar equation to Equation 12 for a specific linguistic response. If we repeat this process for every linguistic response of the two systems, we obtain Equation 12 (for details, see **Supplementary Note 1**). Proceeding in this manner allows us to infer physical laws and new phenomena from an observer in one frame of reference to the observer in the other frame. Next, we will use this process to infer the existence of qualia in ALICE's frame of reference (similar to what Einstein did with the equivalence principle between uniform acceleration and uniform gravitational field). The transformation term $\widetilde{\Lambda}^{\nu\mu}$, is equal to the delta function because of the relativistic principle. Because Alice and ALICE obtain the same measurement results, they are in the same cognitive frame of reference.

So far, we have applied the relativistic principle only to public measurements like reports and phenomenal judgements. But we know that Alice also observes quale Q_{x_t} in her frame of reference. According to broad physicalism, phenomenal consciousness has some kind of physical explanation (physical laws that govern phenomenal consciousness) and as such it is part of physics and part of the physical measurements that Alice can conduct in her first-person frame of reference, ν (even if it's a unique, private measurement). But for the relativistic principle to hold true, all physical laws currently in force in both respective frames of reference should be the same (including the physical laws that enable qualia). Thus, because of our broad physicalist assumption, we can obtain from Equation 11:

$$Q^{\nu}_{x_t} = \widetilde{\Lambda}^{\nu\mu} \widetilde{m}^{\mu} \quad (16)$$

where $\widetilde{m}^{\mu}$ is yet another specific measurement result of ALICE and $Q^{\nu}_{x_t}$ is Alice's quale about input x_t. Just as before, there is a general transformation between the quale that Alice measures in her frame of reference to a specific measurement that ALICE observes about herself in her own frame of reference. Because of the relativistic principle and the equivalence between the two frames of reference, all we need to do is replace ν with μ:

$$Q^{\nu}_{x_t} = \delta^{\nu\mu} \widetilde{m}^{\mu} \quad (17)$$

$$Q^{\mu}_{x_t} = \widetilde{m}^{\mu}. \quad (18)$$

We see that the relativistic principle yields that ALICE will also measure a quale in her frame of reference (and thus zombies cannot exist). We can plug Equation 18 back into Equation 16 and repeat this process for all Alice's qualia. As a result, Alice and ALICE will measure the same qualia about themselves (just like they obtain the same behavioral response and the same phenomenal judgments, assuming we can neglect the small differences due to their differing spatial positions).

This proof leaves the possibility that even though zombies cannot exist because of the relativistic principle, phenomenal consciousness can still be private. To accommodate that, we need to develop yet another mathematical argument, one that is more detailed (eq. 19-35). In addition, the second argument doesn't use the assumption of broad physicalism (at least not explicitly). Notice that until now we haven't used the equations of the cognitive systems. Let's go back to Equation (13), for the linguistic reports:

$$\vec{l}^{\nu}_{p_{l,x_t}} = \widetilde{\Lambda}^{\nu\mu}\vec{l}^{\mu}_{p_{l,x_t}} \quad (19)$$

From equation (10) about the linguistic module, we obtain:

$$L\left(\vec{c}_{p_{l,p_{x_t}}}, E, W, M\right)^{\nu} = \widetilde{\Lambda}^{\nu\mu} L\left(\vec{c}_{p_{l,p_{x_t}}}, E, W, M\right)^{\mu}. \quad (20)$$

We can obtain similar equations for the functional consciousness sub-system (from Equation (9)):

$$\vec{c}^{\nu}_{p_{l,p_{x_t}}} = C\left(\vec{p}_l, \vec{p}_{x_t}, E, W, M, A, I\right)^{\nu} = \widetilde{\Lambda}^{\nu\mu} C\left(\vec{p}_l, \vec{p}_{x_t}, E, W, M, A, I\right)^{\mu} = \widetilde{\Lambda}^{\nu\mu}\vec{c}^{\mu}_{p_{l,p_{x_t}}} \quad (21)$$

Now, consider a special case where Bob asks ALICE and Alice a question about the experience they just had (for example, $\vec{p}_l$ = "What are you experiencing right now?"). In this case, the cognitive system needs to check the phenomenal judgment representation itself. As a result, the functional consciousness subsystem will have the last phenomenal judgments as an input, $\vec{c}_{x_t}$. Plugging it into Equation 9 gives:

$$\vec{c}_{p_{l,c_{x_t}}} = C\left(\vec{p}_l, \vec{c}_{x_t}, E, W, M, A, I\right). \quad (22)$$

The answer, $\vec{c}_{p_{l,c_{x_t}}}$, is yet another phenomenal judgment representation concerning the previous phenomenal judgment and the question. The linguistic module will use this representation to create a linguistic response according to Equation 10:

$$\vec{l}_{p_{l,c_{x_t}}} = L\left(\vec{c}_{p_{l,c_{x_t}}}, E, W, M\right). \quad (23)$$

The linguistic response can be, e.g., $\vec{l}_{p_{l,c_{x_t}}}$ = "I'm experiencing happiness right now." In this scenario we can write the same transformation between ALICE and Alice for the functional consciousness subsystem and the linguistic module as before:

$$\vec{c}^{\nu}_{p_{l,c_{x_t}}} = C\left(\vec{p}_l, \vec{c}_{x_t}, E, W, M, A, I\right)^{\nu} = \widetilde{\Lambda}^{\nu\mu} C\left(\vec{p}_l, \vec{c}_{x_t}, E, W, M, A, I\right)^{\mu} = \widetilde{\Lambda}^{\nu\mu}\vec{c}^{\mu}_{p_{l,c_{x_t}}} \quad (24)$$

$$\vec{l}^{\nu}_{p_{l,c_{x_t}}} = L\left(\vec{c}_{p_{l,c_{x_t}}}, E, W, M\right)^{\nu} = \widetilde{\Lambda}^{\nu\mu} L\left(\vec{c}_{p_{l,c_{x_t}}}, E, W, M\right)^{\mu} = \widetilde{\Lambda}^{\nu\mu}\vec{l}^{\mu}_{p_{l,c_{x_t}}} \quad (25)$$

For all these transformations in Equations 19–21, 24–25, we can substitute the general transformation term, $\widetilde{\Lambda}^{\nu\mu}$, with a delta function, because of the equivalence between Alice's and ALICE's frames of reference (both have the same cognitive system and measurements):

$$\widetilde{\Lambda}^{\nu\mu} = \delta^{\nu\mu} \quad (26)$$

For example, Equations 24–25 will get the form:

$$\vec{c}^{\nu}_{p_{l,c_{x_t}}} = C\left(\vec{p}_l, \vec{c}_{x_t}, E, W, M, A, I\right)^{\nu} = \delta^{\nu\mu} C\left(\vec{p}_l, \vec{c}_{x_t}, E, W, M, A, I\right)^{\mu} = \delta^{\nu\mu}\vec{c}^{\mu}_{p_{l,c_{x_t}}} \quad (27)$$

$$\vec{l}^{\nu}_{p_{l,c_{x_t}}} = L\left(\vec{c}_{p_{l,c_{x_t}}}, E, W, M\right)^{\nu} = \delta^{\nu\mu} L\left(\vec{c}_{p_{l,c_{x_t}}}, E, W, M\right)^{\mu} = \delta^{\nu\mu}\vec{l}^{\mu}_{p_{l,c_{x_t}}} \quad (28)$$

Now, let's switch to Alice's frame of reference. From Alice's point of view, she has phenomenal consciousness. This perspective is also a physical frame of reference. From her first-person perspective, she experiences happiness (quale Q_{x_t}) and a question, Q_l = "What are you experiencing right now?" After the question, she will have an experience of an answer to the question, $Q_{Q_l,Q_{x_t}}$, and accordingly she will respond with $\vec{l}_{Q_l,Q_{x_t}}$ = "I'm experiencing happiness right now." Notice that from her first-person frame of reference, she directly experiences her qualia and uses the quale (e.g., of happiness) to answer the question (in her head) and to articulate the answer as a linguistic report. We can formulate an equation according to her first-person frame of reference:

$$\vec{l}^{\nu}_{Q_l,Q_{x_t}} = F\left(Q_{Q_l,Q_{x_t}}\right)^{\nu}, \quad (29)$$

where $Q_{Q_l,Q_{x_t}}$ is the quale of the answer, $\vec{l}_{Q_l,Q_{x_t}}$ is the linguistic response, and F is a general function mapping a quale to a vector of the linguistic response. From Alice's frame of reference, she experiences herself trying to answer the question in her head $\left(Q_{Q_l,Q_{x_t}}\right)$ and then responds accordingly $(\vec{l}_{Q_l,Q_{x_t}})$. Her direct experience is that her linguistic response, $\vec{l}_{Q_l,Q_{x_t}}$ is a result of the quale of her answer, $Q_{Q_l,Q_{x_t}}$.

We can now make use of the equivalence between Alice and ALICE. First, notice that we can describe Alice's linguistic response according to her first-person perspective (Equation 29) and according to a third-person perspective (Equation 23). Both descriptions refer to the same linguistic response, and thus $\vec{l}^{\nu}{}_{Q_l,Q_{x_t}} = \vec{l}^{\nu}_{p_{l,c_{x_t}}}$. Plugging this result into Equation 29 gives us:

$$\vec{l}^{\nu}_{p_{l,c_{x_t}}} = F\left(Q_{Q_l,Q_{x_t}}\right)^{\nu}. \tag{30}$$

According to Equation 28, because of the equivalence between Alice's and ALICE's frames of reference for linguistic responses, we can replace the left-hand side of Equation 30 with:

$$\delta^{\nu\mu}\vec{l}^{\nu}_{p_{l,c_{x_t}}} = F\left(Q_{Q_l,Q_{x_t}}\right)^{\nu}. \tag{31}$$

Because of the delta function, we know that $\nu = \mu$, and we can substitute μ with ν on the right-hand side of Equation 31:

$$\vec{l}^{\nu}_{p_{l,c_{x_t}}} = F\left(Q_{Q_l,Q_{x_t}}\right)^{\mu}. \tag{32}$$

We have thus completed the transformation to ALICE's frame of reference. ALICE, like Alice, must have qualia from her first-person perspective, which she uses to create a linguistic response. As a result of Equation 32, we notice again that the relativistic principle yields that ALICE, as cognitive frame of reference μ, will also observe a quale in her frame of reference. This invalidates Chalmers' (1996) assumption that a zombie can be physically and computationally isomorphic to a phenomenally conscious person (which we assumed in the beginning by stating that ALICE was a zombie). Notice that now we can use Equation 23 and substitute $\vec{l}^{\mu}_{p_{l,c_{x_t}}}$ in Equation 32 with:

$$L\left(\vec{c}_{p_{l,c_{x_t}}}, E, W, M\right)^{\mu} = F\left(Q_{Q_l,Q_{x_t}}\right)^{\mu}. \tag{33}$$

There is an equivalence between the form of the right-hand side of the equation and the form of the left-hand side of the equation (in both sides there is a function that gets input). We can identify the function F as the function of the linguistic module, $F = L$, and the input of function F as the input of function L (notice that the modules E, W, M are not the inputs of L, they are just modules that the dynamics of the language module uses to create the linguistic response according to the input $\vec{c}_{p_{l,c_{x_t}}}$):

$$\vec{c}^{\mu}_{p_{l,c_{x_t}}} = Q^{\mu}_{Q_l,Q_{x_t}}. \tag{34}$$

We here get an interesting result in ALICE's frame of reference. Not only will ALICE observe a quale just like Alice, but this quale is exactly her complex integrated representation that bears content about phenomenology ('phenomenal judgement'). We can identify between the inputs $\vec{c}^{\mu}_{p_{l,c_{x_t}}}$ and $Q^{\mu}_{Q_l,Q_{x_t}}$ because in conscious agents like humans there is a close relation between the two. In fact, there is an isomorphism between them, and we can do a one-to-one correspondence mapping between them. As we saw, every quale will always be accompanied with a corresponding phenomenal judgment and every phenomenal judgment will always be accompanied with a corresponding quale. This pair of quale and phenomenal judgment is unique for each input $\vec{x}_t$ and every time $\vec{x}_t$ is present for the conscious cognitive system, if one part of the pair is present then the other part will be present as well. We saw that zombies cannot exist because they must have qualia according to the relativistic principle (eq. 32). We also saw that in the cognitive frame of reference that uses phenomenal judgments there is an equality between the functions that use qualia and phenomenal judgments (eq. 33). The identity between the functions and the isomorphism between qualia and phenomenal judgments are enough to show that in the cognitive frame of reference that uses phenomenal judgments there is an equality between qualia and phenomenal judgments (eq. 34). We reach an equivalence between phenomenal consciousness and phenomenal judgments that were created by cognitive systems, and hence obtain an equivalence between functional consciousness and phenomenal consciousness. It is easy to show that if we use the equivalence principle once again to move back to Alice's frame of reference, we will obtain the same result for her, as well. We developed this equation from a special case of a quale about a quale (the experience of an experience, for example, $\vec{l}_{Q_l,Q_{x_t}} =$ "I'm experiencing happiness right now."), but we can repeat the same process for different cases of qualia, such as a quale of an arbitrary input (seeing an apple for example), Q_{x_t}. Generally, from the first-person perspective, a quale causes something to happen, whether it will cause a linguistic response (like in Equation 29) or the next associative thought. Because there is always an equivalent phenomenal judgment representation that will do the same thing in the cognitive system, we can always repeat a similar process like we did here (Equations 29–34) with an adequate equation similar to Equation 29 for the specific case. Furthermore, we can use the identity in Equation 34 in these similar equations to prove similar identities for different kinds of qualia (see **Supplementary Note 2** for details). The relativistic principle ensures that because we cannot distinguish between the quale and the phenomenal judgment representation that accompanies it, they are the same thing. Consequently, we can generalize Equation 34:

$$\vec{c}^{\mu}_{x_t} = \vec{Q}^{\mu}_{x_t} \tag{35}$$

In the first-person frame of reference, there is an equality between the quale and the phenomenal judgment-representation, and hence the quale is a point in an appropriate state space and can be represented as a vector. The first-person frame of reference is the frame of the cognitive system measuring itself. In other words, in its dynamics, the cognitive system uses its own representations as inputs and outputs and they interact with each other (like in our cognitive system model, Equations 1–10).

Conditions for Third- and First-Person Frames of Reference

What happens when we move to different cognitive frames of reference? Why can't they directly measure the quale of the other frame of reference? Why is there a difference between first-person and third-person perspectives, if it's not because of some phenomenal property of privacy? To answer these questions, let's expand the equations beyond the simple case of the equivalence between a conscious human and a copy of the human's cognitive system (a zombie-like system). We showed that in the simple case of Alice and ALICE, not only are both cognitive systems equivalent, but they are also in the same cognitive frame of reference, $\nu = \mu$ (that's why we could use the delta function).

Now, let's examine two cognitive frames of reference that are not a copy of each other: Alice and Bob. These two cognitive systems developed separately and there is no reason to assume that they learned to associate inputs with the same neural patterns. Each associated different states as representations from the state space according to its own personal developmental history. As a result, Alice and Bob have no states in common and the set of stats that Alice's cognitive system uses is disjoint with the set of states that Bob's cognitive system uses. As before, from her first-person frame of reference, Alice has an experience of an apple (quale $\vec{Q}^{\nu}_{x_t}$). What will happen if Bob tries to directly observe the quale of Alice in his cognitive frame of reference? We can describe this process from his cognitive frame of reference, μ. His perception, functional consciousness, and linguistic subsystems (Equations 6–8) directly obtain Alice's quale:

$$P\left(\vec{Q}^{\nu}_{x_t}, E, W, M\right)^{\mu} = P\left(\vec{c}^{\nu}_{x_t}, E, W, M\right)^{\mu} = \emptyset \quad (36)$$

$$C\left(\vec{Q}^{\nu}_{x_t}, E, W, M, A, I\right)^{\mu} = C\left(\vec{c}^{\nu}_{x_t}, E, W, M, A, I\right)^{\mu} = \emptyset \quad (37)$$

$$L\left(\vec{Q}^{\nu}_{x_t}, E, W, M\right)^{\mu} = L\left(\vec{c}^{\nu}_{x_t}, E, W, M\right)^{\mu} = \emptyset, \quad (38)$$

where Ø is the empty set. To measure Alice's quale directly means to use it directly in Bob's cognitive frame of reference. Because her quale is equal to a phenomenal judgment representation (Equation 35), we substitute $\vec{Q}^{\nu}_{x_t}$ with $\vec{c}^{\nu}_{x_t}$. But because Bob and Alice are two different cognitive frames of reference, Bob's cognitive system doesn't recognize Alice's representation, and we get an empty set. As a result, Bob cannot directly measure Alice's quale. The only solution for Bob is to indirectly measure it using his sensation subsystem (Equation 3).

What will Bob measure while using his sensation subsystem to measure Alice's quale? Recall that the sensation subsystem measures physical properties, y_j from outside of the cognitive system, that can be detected by its sensors. According to eq. 35, qualia, like phenomenal judgments, are neural patterns or states of the state space of Alice's cognitive system (that are being measured directly in Alice's cognitive frame of reference). every degree of freedom (variable) of the dynamics of the system is an axis of this space (e.g., the axes can be firing rate, membrane potential, position of firing neurons, etc.). The axes form the basis vectors of the state space, $\sum_{m=1}^{N} \hat{e}_m = \hat{e}_1 + \hat{e}_2 + \ldots \hat{e}_N$ where $\sum \hat{e}$ means a summation over the basis vectors and m = 1,2,3. . .N is an index that runs from 1 to N, according to the N-dimensions of the state space. Each state can be written, in a unique way, as $\sum_{m=1}^{N} \alpha_m \hat{e}_m$, where $\alpha_1 \ldots \alpha_N$ are coefficients. These Numbers are the coordinates of the axes (e.g., suppose that for the neural representation of an apple there is a specific firing rate. This is the coordinate α_1 of the first axis $\hat{e}_1$, the axis of firing rates and we will denote it by, $\alpha_1 \hat{e}_1$.) These variables and coefficients of the dynamics of Alice's cognitive system are the physical properties y_j that Bob's sensation subsystem will measure in the case of Alice's quale/representation, $y_j = \sum_{m=1}^{j} \alpha_m \hat{e}_m$ (e.g., without losing generality, let's assume that Alice's quale/phenomenal judgement while seeing an apple will have a specific pattern of 3 variable: firing rates, membrane potentials and position of firing neurons. Each of them with a specific coefficient. Bob's sensation subsystem will measure these 3 properties with their specific coefficients.):

$$T\left(\vec{Q}^{\nu}_{x_t}\right)^{\mu} = T\left(\vec{c}^{\nu}_{x_t}\right)^{\mu} = \vec{t^i}^{\mu}_{Q_{x_t}} = \vec{t^i}^{\mu}_{\sum \alpha_m \hat{e}_m}, \quad (39)$$

where T^{μ} is Bob's sensation subsystem, $\vec{Q}^{\nu}_{x_t}$ is Alice's quale, $\vec{c}^{\nu}_{x_t}$ is Alice's phenomenal judgment representation, $\vec{t^i}^{\mu}_{Q_{x_t}}$ are the outputs of the sensation subsystem of Bob about Alice's quale. $\sum \alpha_m \hat{e}_m$ are the variables and coefficients of Alice's quale/representation, and $\vec{t^i}^{\mu}_{\sum \alpha_m \hat{e}_m}$ are the outputs of the sensation subsystem of Bob about the physical properties of Alice's quale. Bob's sensation subsystem is the only part of the cognitive system that can measure Alice's quale, though indirectly. As before, we substitute her quale with the corresponding representation. The sensation subsystem can measure the physical properties of the representation and create an output that Bob's cognitive frame of reference can utilize (e.g., a visual sensation of a brain pattern). Only then will Bob's cognitive frame create its own percept, and eventually a quale of this sensation. Notice that this quale is about the indirect measurement of Alice's quale, a brain pattern that Bob measured from Alice's brain:

$$C\left(P\left(\vec{t^i}_{\sum \alpha_m \hat{e}_m}\right)\right)^{\mu} = \vec{c}^{\mu}_{\sum \alpha_m \hat{e}_m} = \vec{Q}^{\mu}_{\sum \alpha_m \hat{e}_m} \quad (40)$$

Notice the difference between Alice's initial quale, $\vec{Q}^{\nu}_{x_t}$, and Bob's final quale about her brain pattern, $\vec{Q}^{\mu}_{\sum \alpha_m \hat{e}_m}$. Interestingly, even in the same cognitive frame of reference (of Alice, for example), if we introduce to Alice's sensation subsystem one of Alice's qualia, the sensation subsystem will still not recognize it directly as a quale. Instead, the subsystem will measure the physical properties corresponding to that quale, $y_j = \sum_{m=1}^{j} \alpha_m \hat{e}_m$. For example, we can show Alice an activation pattern going on in her brain as she reports seeing an apple. Alice will not recognize this pattern as her representation of seeing an apple and will instead see a visual image of the brain pattern. This is because the sensation subsystem still doesn't use representations, let alone the representations of higher cognitive subsystems like perception and cognition. It just measures physical properties from outside of the cognitive system:

$$T\left(\vec{Q}^{\nu}_{x_t}\right)^{\nu} = T\left(\vec{c}^{\nu}_{x_t}\right)^{\nu} = \vec{t^i}^{\nu}_{\sum \alpha_m \hat{e}_m} \quad (41)$$

Only the parts of the cognitive system that use the phenomenal judgment representation, $\vec{c}^{\nu}_{x_t}$, will measure it directly (such as the functional consciousness subsystem and the language module). The relativistic principle ensures us that this direct measurement of a phenomenal judgment representation is a quale (Equation 35). Any other cognitive frames of reference will just measure the physical substance of the representation *via* the sensation subsystem. It seems that the reason we have such a distinct difference between first-person and third-person perspectives is due to the direct measurement of the phenomenal judgment representation in the functional consciousness subsystem, and the function of the sensation subsystem that specializes in measuring bare physical properties and not representations. In sum, the condition for the ability to measure a quale (viz., the first-person frame of reference) is to have a subsystem that uses the corresponding phenomenal judgment representation (i.e., to measure directly phenomenal judgment. Because then the functional consciousness subsystem uses the phenomenal judgment in its dynamics and according to eq. 35, we can substitute it with the appropriate quale):

$$C\left(\vec{c}^{\mu}_{x_t}, E, W, M, A, I\right)^{\mu} = C\left(\vec{Q}^{\mu}_{x_t}, E, W, M, A, I\right)^{\mu} = \vec{c}^{\mu}_{c_{x_t}} = \vec{Q}^{\mu}_{Q_{x_t}} \quad (42)$$

Here, the functional consciousness subsystem, C^{μ}, uses in its dynamics (directly measures) a phenomenal judgment representation, $\vec{c}^{\mu}_{x_t}$, and hence measures it directly to create a new phenomenal judgment representation about it, $\vec{c}^{\mu}_{c_{x_t}}$. According to the relativistic principle, these phenomenal judgments are measured as qualia, $\vec{Q}^{\mu}_{x_t}$, $\vec{Q}^{\mu}_{Q_{x_t}}$. The condition to have a third-person frame of reference is to activate the sensation subsystem (39, 41):

$$T\left(\vec{Q}_{x_t}\right)^{\mu} = \vec{ti}^{\mu}_{\sum \alpha_m \hat{e}_m} \quad (43)$$

Notice that the cognitive frame of reference that measured the quale is irrelevant for the sensation subsystem. The output will always be the same: sensations of the physical properties of the quale/phenomenal judgement, $\vec{ti}^{\mu}_{\sum \alpha_m \hat{e}_m}$.

Transformation Between Cognitive Frames of Reference and Between First- and Third-Person Frames of Reference

Finally, according to the principle of relativity, all admissible frames of reference measure the same laws of physics that are currently in force, and hence the equations describing the laws of physics have the same form in all admissible frames of reference (like Equation 1, describing the Lorentz transformation between all inertial frames of reference). As a relativistic phenomenon, consciousness should also have a similar transformation that preserves the same form of the transformation equation for all admissible cognitive frames of reference. Let's develop this equation. The transformation should take us from the measurements of one frame, μ (Bob), to the measurements of another frame, ν (Alice). Regardless of the measurements, the equation should look the same (for example, if one of the measurements causes new terms to appear in the equation, the form of the equation is no longer the same, and thus the equation is not relativistic because we can distinguish between the measurements according of their different terms. As a result, the equivalence between the admissible cognitive frames of reference has broken. We need to ensure that this is not the case in our equation). Notice that, in contrast to the scenario of Alice and ALICE, now the transformation has no constraints in the form of a question that guides the observers. Instead, they measure the same input, and the transformation should give the answer for what the cognitive system in each frame of reference measures. When we move from measurements of a quale in one cognitive frame of reference, μ, to measurements of a quale in a second cognitive frame of reference, ν, there are three cases of what they can measure. The cognitive frames can measure a third input (e.g., an apple), they can measure frame μ, or they can measure frame ν. If they measure a third input, like an apple, each of the frames will have its own quale about the apple, $\vec{Q}^{\nu}_{x_t}$, $\vec{Q}^{\mu}_{x_t}$. If they measure the quale of frame μ, then frame μ measures the quale $\vec{Q}^{\mu}_{Q_{x_t}}$ directly, while frame ν will measure the physical properties of the quale of frame μ, $\vec{Q}^{\nu}_{\sum \alpha_m \hat{e}_m}$, and vice versa if they measure the quale of frame ν.

We need to use another transformation term, $\Lambda^{\nu\mu}$, for this general scenario. It needs to give the correct measurements for each case, while preserving the form of the transformation equation. Because we lack the accurate equations of the cognitive system and use general forms in the equations, the transformation equation will also be in a general form:

$$\vec{Q}^{\nu}_{x_t} = \Lambda^{\nu\mu}\left(\vec{Q}^{\mu}_{x_t}\right), \quad (44)$$

where $\Lambda^{\nu\mu}$ is the transformation function from cognitive frame μ to cognitive frame ν, and $\vec{Q}^{\nu}_{x_t}$, $\vec{Q}^{\mu}_{x_t}$ are the qualia of input x_t as measured in frames ν, μ respectively.

In order to build $\Lambda^{\nu\mu}$, for the transformation to operate successfully, one key element is to check the physical properties, y_j of the input. To that end we introduce the function m_C that checks if the physical properties y_j of the input $\vec{x}_t$ are the physical properties of the representations that the functional consciousness subsystem uses as outputs, $\sum_{m=1}^{j} \alpha_m \hat{e}_m$. In other words, it checks whether the input is a state of the state space of the functional consciousness subsystem C (i.e., phenomenal judgments). If it does then m_C returns the state $\sum_{m=1}^{j} \alpha_m \hat{e}_m$ and if the input is not part of the state space of C, then m_C returns the empty set Ø:

$$\begin{aligned} m_C\left(\vec{x}_t\right) &= \sum_{m=1}^{j} \alpha_m \hat{e}_m \leftrightarrow C\left(\vec{x}_t, E, W, M, A, I\right) \\ &= C\left(\sum_{m=1}^{j} \alpha_m \hat{e}_m, E, W, M, A, I\right) \\ &= C\left(\vec{c}_{\tilde{x}_t}, E, W, M, A, I\right) \\ & \textit{else } m_C(\vec{x}_t) = \emptyset. \end{aligned} \quad (45)$$

If the input $\vec{x}_t$ is a state of the state space of C then it's a phenomenal judgment, $\vec{x}_t = \sum_{m=1}^{j} \alpha_m \hat{e}_m = \vec{c}_{\tilde{x}_t}$ where $\vec{c}_{\tilde{x}_t}$ is a phenomenal judgment about some input $\vec{\tilde{x}}_t$. In that case, m_C returns this state. Notice that because C can get phenomenal judgments as inputs (see eq. 22 and 42) $\vec{x}_t$ can be input of C and thus can be measured directly by the functional consciousness subsystem. If the input is not part of the state space of C, then m_C returns Ø. Now we can introduce the transformation function

$\Lambda^{\nu\mu}$:

$$\Lambda^{\nu\mu}\left(\vec{x}_t, S^{\nu}\right) = \begin{cases} m_C^{\nu}\left(\vec{x}_t\right) = \emptyset \vee \delta^{\vec{r}_0^{\nu}\vec{r}_{x_t}^{\nu}} = 0, & S^{\nu}\left(\vec{x}_t\right) \\ m_C^{\nu}\left(\vec{x}_t\right) \neq \emptyset \wedge \delta^{\vec{r}_0^{\nu}\vec{r}_{x_t}^{\nu}} = 1, & C^{\nu}\left(\vec{x}_t\right) \end{cases} \quad (46)$$

Where $\vee$ is a symbol of the logic operation 'or' and $\wedge$ is a symbol of the logic operation 'and.' In order to carry out the transformation from cognitive frame μ to cognitive frame ν, the transformation function $\Lambda^{\nu\mu}$ gets as inputs the input $\vec{x}_t$ and cognitive system of frame ν, S^{ν}. $\Lambda^{\nu\mu}$ checks two conditions, whether $\vec{x}_t$ is a state of the state space of C^{ν} (the functional consciousness subsystem of frame ν), and whether $\vec{x}_t$ is in $\vec{r}_0^{\nu}$, the spatial position of cognitive system S^{ν} (as seen from its own cognitive frame of reference, ν). The spatial position of $\vec{x}_t$ as seen from cognitive frame of reference ν, is $\vec{r}_{x_t}^{\nu}$. Spatial positions $\vec{r}_0^{\nu}$, $\vec{r}_{x_t}^{\nu}$ are part of the physical properties of the cognitive system of frame ν and the input (respectively). The transformation function checks the first condition by plugging $\vec{x}_t$ into m_C^{ν} (that checks whether the input is a state of the state space of C in the cognitive frame of reference ν) and checks the second condition by using a delta function, $\delta^{\vec{r}_0^{\nu}\vec{r}_{x_t}^{\nu}}$, that equals 1 if $\vec{r}_{x_t}^{\nu} = \vec{r}_0^{\nu}$ and 0 if $\vec{r}_{x_t}^{\nu} \neq \vec{r}_0^{\nu}$. If $\vec{x}_t$ is not a state in the space state of C^{ν}, then m_C^{ν} will return the empty set. This means that the input $\vec{x}_t$ cannot be measured directly by the functional consciousness subsystem of frame ν. The same situation occurs if the delta function is equal to 0, it means that $\vec{x}_t$ is not in the cognitive system's spatial position and thus is not available for C^{ν} to use it and measure it directly. As a result, $\vec{x}_t$ is sent as an input to cognitive system S^{ν} for an indirect measure. There, the sensation subsystem of frame ν gets $\vec{x}_t$ as input. From there, the process will continue along the hierarchy of the cognitive system to the perception subsystem and the functional consciousness subsystem of frame ν (see eq. 2, 5). If $\vec{x}_t$ is a state in the state space of C^{ν}, then m_C^{ν} will not return the empty set. This means that the input can be measured directly by the functional consciousness subsystem of frame ν. In addition, the second condition is being checked and if the delta function is equal to 1, it means that the input is in the spatial position of the cognitive system and thus available for C^{ν}. As a result, if the two conditions are fulfilled, $\vec{x}_t$ is sent as an input directly to C^{ν}, the functional consciousness subsystem of cognitive system S of frame ν, to measure the input directly as a quale. In sum, cognitive system S of frame ν is in the third-person perspective if either $\delta^{\vec{r}_0^{\nu}\vec{r}_{x_t}^{\nu}} = 0$ or m_C^{ν} $(\vec{x}_t) = \emptyset$ because it uses the sensation subsystem (condition for third-person frame of reference is fulfilled, eq. 43) and cognitive system S of frame ν is in the first-person perspective only if both $\delta^{\vec{r}_0^{\nu}\vec{r}_{x_t}^{\nu}} = 1$ and m_C^{ν} $(\vec{x}_t) \neq \emptyset$, because then it uses C^{ν} $(\vec{x}_t)$ which measures directly the phenomenal judgment (i.e., a quale. Condition for first-person frame of reference is fulfilled, eq. 42).

Let's check if Equation 46 gives the correct measurements for each case of the transformation from frame μ to ν while preserving the form of the transformation equation (Equation 44). We mentioned three cases: the cognitive frames can measure a third input, (e.g., apple), they can measure frame μ, or they can measure frame ν. In the first case, we start from a quale of frame μ about a third input like an apple, $\vec{Q}_{x_t}^{\mu}$. The transformation function checks if the input $\vec{x}_t$ is a state from the state space of C^{ν} or is it inside the cognitive system S^{ν}. Because this is an input from outside of the functional consciousness subsystem and the cognitive system, $\vec{x}_t$ (the apple) is not a state of the subsystem and not in the spatial position of the cognitive system, and hence the transformation function sends $\vec{x}_t$ to S^{ν}, the cognitive system in frame ν. As a result, S^{ν} will use its sensation, perception, and functional consciousness subsystems to create its own quale of the apple, $\vec{Q}_{x_t}^{\nu}$. The transformation function succeeds in giving the correct output for the first case. In the second case, frame μ measures its own quale $\vec{x}_t = \vec{Q}_{\tilde{x}_t}^{\mu}$ and thus in the transformation equation (eq. 44), $\vec{Q}_{x_t}^{\mu} = \vec{Q}_{Q_{\tilde{x}_t}}^{\mu}$. The transformation function checks if the input $\vec{x}_t$ is a state from the state space of C^{ν} and if it is part of the cognitive system S^{ν}. Because this is again an input from outside of the cognitive system and the functional consciousness subsystem of frame ν (it's a quale from frame μ), the transformation function sends the input to S^{ν}. As a result, frame ν will use its sensation, perception, and functional consciousness subsystems to create a quale of the physical properties of the quale from frame μ, and thus in the transformation equation (eq. 44) $\vec{Q}_{x_t}^{\nu} = \vec{Q}_{\sum \alpha_m \hat{e}_m}^{\nu}$. The transformation function succeeds in giving the correct output for the second case as well. In the last case, frame μ measures a quale $\vec{x}_t = \vec{Q}_{\tilde{x}_t}^{\nu}$ of frame ν. As a result, frame μ will measure the physical properties of the quale from frame ν, and thus in the transformation equation (eq. 44), $\vec{Q}_{x_t}^{\mu} = \vec{Q}_{\sum \alpha_m \hat{e}_m}^{\mu}$. The transformation function checks if the input $\vec{x}_t$ is a state from the state space of C^{ν} and if it's part of the cognitive system S^{ν}. This time, the answer is positive for both conditions. Because this is an input from within the functional consciousness subsystem of frame ν, the matching function of the subsystem m_C^{ν} will return something other than the empty set and the delta function will be equal to 1. Consequently, the transformation function sends $\vec{x}_t$ directly to the functional consciousness subsystem of frame ν [where, $C^{\nu}(\vec{x}_t) = C^{\nu}(\sum \alpha_m \hat{e}_m) = C^{\nu}(\vec{c}_{\tilde{x}_t}) = \vec{c}_{c_{\tilde{x}_t}}^{\nu}$, see Equation 45]. As a result, frame ν will measure its own quale, and thus in the transformation equation (eq. 44) $\vec{Q}_{x_t}^{\nu} = \vec{Q}_{Q_{\tilde{x}_t}}^{\nu}$. The transformation function also succeeds in giving the correct output for the last case. Notice that the transformation Equation 44 maintains the same form regardless of the specific measurement case or which cognitive frames of reference are involved. Hence, Equation 44 satisfies the relativity principle as desired.

All the preceding scenarios can give the impression that the two conditions are redundant. They always seem to agree with each other and so, one condition is enough to get the correct transformation. But now let's choose a special case where $\nu = \mu$. Hence, the transformation function will be $\Lambda^{\nu\nu}(\vec{x}_t, S^{\nu}) = \Lambda^{\mu\mu}(\vec{x}_t, S^{\mu})$. There are two scenarios for this case. One scenario is the identity transformation in which the transformation leaves us in the same cognitive system that we started from (and thus in the same cognitive frame of reference). According to Equation 46, $\Lambda^{\mu\mu}$ gives the correct answer in this scenario, because we stay in the same cognitive system all the time (same frame and same spatial location). In other words, the transformation function doesn't change anything, $\Lambda^{\mu\mu} = 1 \Leftrightarrow \Lambda^{\nu\mu} = \delta^{\nu\mu}$. We can call this the full symmetry scenario, because the two cognitive frames are completely symmetric and we cannot distinguish between them (in other words, we started from a cognitive system and come back after the transformation to the same system). The transformation Equation 44 will be in this scenario,

$$\vec{Q}^{\nu}_{x_t} = \delta^{\nu\mu}(\vec{Q}^{\mu}_{x_t}). \tag{47}$$

The second scenario, however, is more complicated. Here we have two cognitive systems but they are in the same cognitive frame of reference. They have the same dynamics and thus the same representations and the same qualia, but they remain two discrete systems with different spatial positions. As a result, for the same input both will have the same answer to every question they might be asked. This is the scenario that we started from when we developed the equivalence principle between ALICE and Alice. It's very special scenario where we have two cognitive systems in different spatial positions but in the same cognitive frame. Notice, however, that even though they are in the same cognitive frame, because they have different spatial positions, they cannot measure directly the representations of the other system and thus they must use their sensation subsystem and measure indirectly the quale of the other system (similar to Equation 41 where the same cognitive system will still measure the physical properties of its own quale due to the sensation subsystem). In other words, although they are in the same cognitive frame of reference, Alice will measure ALICE from a third-person perspective and vice versa. Will Equation 46 give the correct transformation from frame μ (ALICE) to ν (Alice) also in this scenario?

For the first case, both Alice and ALICE measure a third input like an apple. In this case, when we transfer from ALICE to Alice, we will get an answer just like in the previous, full symmetric scenario. Because the physical properties of the apple is from the outside of the cognitive system S^{ν}, the transformation function will use the sensation, perception, and functional consciousness subsystems of the system in frame ν to create the appropriate quale, $\vec{Q}^{\nu}_{x_t}$. But because both systems are in the same cognitive frame, their qualia are the same and the quale of the input will remain the same. As a result, the transformation gives us the correct answer for this case,

$$\vec{Q}^{\nu}_{x_t*} = \delta^{\nu\mu}(\vec{Q}^{\mu}_{x_t*}), \tag{48}$$

where $\vec{x}_t{}^*$ is input from outside of both cognitive frames. ALICE and Alice will measure the same quale about an input that is outside of both their cognitive frames.

If ALICE (μ) measures its own quale, $\vec{Q}^{\mu}_{x_t} = \vec{Q}^{\mu}_{Q_{\tilde{x}_t}}$, the transformation function checks if the input $\vec{x}_t = \vec{Q}^{\mu}_{\tilde{x}_t}$ is a state from the state space of C^{ν} and if it's in the position of the cognitive system S^{ν}. In this case, because $\mu = \nu$, $\vec{x}_t$ is a state of the functional consciousness subsystem of frame ν as well (the state $\vec{x}_t = \sum_{m=1}^{j} \alpha_m \hat{e}_m$ is the same state in both frames). But the other condition fails because the systems are not in the same spatial position, and $\vec{x}_t$ comes from the position of the other cognitive system, S^{μ}. As a result, $m_C{}^{\nu}(\vec{x}_t) \neq \emptyset$, but $\delta^{\vec{r}_0{}^{\nu}\vec{r}_{x_t}{}^{\nu}} = 0$. Because only one condition is fulfilled for the first-person perspective and not both, the transformation function will send $\vec{x}_t$ to S^{ν} for an indirect measure. From the sensation subsystem in frame ν until eventually creating a quale of the physical properties of the quale from frame μ, $\vec{Q}^{\nu}_{x_t} = \vec{Q}^{\nu}_{\sum \alpha_m \hat{e}_m}$. The transformation function succeeds in giving the correct answer also for this case. If ALICE (μ) tries to measure the quale of Alice (ν) the result will be a quale of the physical properties of Alice's quale, $\vec{Q}^{\mu}_{x_t} = \vec{Q}^{\mu}_{\sum \alpha_m \hat{e}_m}$. The transformation function checks if the physical properties of the input $\vec{x}_t = \sum_{m=1}^{j} \alpha_m \hat{e}_m$ is a state from the state space of C^{ν} and if it's in the position of the cognitive system S^{ν} of frame ν (Alice). This time, the answer is positive for both conditions. Because the input is also a state from the state space of the functional consciousness subsystem of frame ν and it is in the position of S^{ν}, the matching function of the subsystem m^{ν}_C will return something other than the empty set and the delta function will be equal to 1. Consequently, the transformation function sends $\vec{x}_t$ directly to the functional consciousness subsystem of the cognitive system in frame ν. As a result, frame ν will measure its own quale, $\vec{Q}^{\nu}_{x_t} = \vec{Q}^{\nu}_{Q_{\tilde{x}_t}}$. The transformation function also succeeds in giving the correct output for the last case. Now we see that the last two cases in the scenario of ALICE and Alice are different than the full symmetry scenario. The different spatial positions of the cognitive systems break the symmetry of the identity transformation. Although the systems are in the same cognitive frame, the different positions cause the transformation to be different than $\delta^{\nu\mu}$. The different spatial positions break the symmetry because they cause the cognitive systems to use their sensation subsystem to measure the quale of the other system. We can call this scenario the broken symmetry scenario. As a result, we can write the transformation Equation 44 for this scenario as:

$$\vec{Q}^{\nu_{r_0}}_{x_t**} = \delta^{\nu\mu}\Lambda^{\nu_{r_0}\,\mu_{r'_0}}\left(\vec{Q}^{\mu_{r'_0}}_{x_t**}\right), \tag{49}$$

Which gives us -

$$\vec{Q}^{\nu_{r_0}}_{x_t**} = \Lambda^{\nu_{r_0}\,\nu_{r'_0}}\left(\vec{Q}^{\nu_{r'_0}}_{x_t**}\right). \tag{50}$$

Here $\vec{x}_t{}^{**}$ is an input of a state from the state space of C either from frame ν or μ in the broken symmetry case. The delta function is presented because both cognitive systems are in the same frame, but $\Lambda^{\nu_{r_0}\,\nu_{r'_0}}$ is still different than the identity transformation because the symmetry of the cognitive systems was broken by their different spatial positions, r_0, r'_0 (for another example of the importance of the spatial positions condition, see **Supplementary Note 3**). Because of the broken symmetry scenario, we define the first-person frame of reference slightly differently than the definition of a cognitive frame of reference. The first-person frame of reference is defined not only by the dynamics of the cognitive system but also by its position. As a result, Both Alice and ALICE will have the same answers to the question of what is it like (because they are in the same cognitive frame of reference), but because they are not in the same spatial position, they are not in the same first-person frame of reference. When one is measuring the dynamics of the other, all they see is the physical patterns of the other cognitive system. In other words, when one cognitive frame is in another spatial position, it has to use its sensation subsystem and thus satisfy the condition for the third-person frame of reference (eq. 43). Only if the two identical cognitive systems can be in the same position can they share the same dynamics and use their functional consciousness subsystem directly without the need to use the sensation subsystem. Only then will they satisfy the condition of the first-person frame of reference (eq. 42).

Now we also see that the general transformation, $\widetilde{\Lambda}^{\nu\mu}$, that we used for the equivalence principle in order to prove eq. 35 is a special case of the transformation function we developed here, $\Lambda^{\nu\mu}$. In this case, the two cognitive frames of reference are the same but have different spatial positions (e.g., Alice and ALICE), and their input is constrained to be a question (the same question for both cognitive frames of reference). Because both frames are the same and because their input is the same question, we always get a scenario like eq. 48 and thus, $\widetilde{\Lambda}^{\nu\mu} = \delta^{\nu\mu}$.

The conclusion is that the transformation function $\Lambda^{\nu\mu}$ gives the correct answers to all cases and scenarios while preserving the form of the transformation equation (equation 44) as desired. Equation 46 for $\Lambda^{\nu\mu}$ can be approximated by a simpler form:

$$\Lambda^{\nu\mu}\left(\vec{x}_t, S^\nu\right) = \left(1 - \delta^{\vec{r}_0^\nu \vec{r}_{x_t}^\nu}\right) S^\nu(\vec{x}_t) + \delta^{\vec{r}_0^\nu \vec{r}_{x_t}^\nu}\, C^\nu(\vec{x}_t) \qquad (51)$$

The first term is the condition to be in a third person perspective because of the use of the sensation subsystem. Only when $\delta^{\vec{r}_0^\nu \vec{r}_{x_t}^\nu} = 0$ will this term be equal to 1. The second term is the condition of being in the first-person perspective (because of the use of the functional consciousness subsystem). Only when $\delta^{\vec{r}_0^\nu \vec{r}_{x_t}^\nu} = 1$ will this term be equal to 1. This approximation checks only if the spatial position of the input is the same as the spatial position of the cognitive system and it gives the correct answers to all scenarios. It transforms from the third-person frame of reference to the first-person frame of reference and vice versa. But it is only an approximation because it does not take into account the matching function $m_C(\vec{x}_t)$ and so in the case of two cognitive systems from the same cognitive frames of reference (like Alice and ALICE) the approximation will neglect this information and the result will be two different cognitive frames ν, μ instead of just one.

In **Supplementary Note 4**, we show the inverse transformation function, from frame of reference ν to frame μ and that the transformation function and the inverse transformation function cancel each other out (giving the identity transformation), as expected.

DISCUSSION

Consciousness as a Relativistic Phenomenon

As we saw, the assumption that the relativity principle also includes measurements of cognitive systems forces us to treat consciousness as a relativistic phenomenon. If the relativity principle is true, then zombies are not nomologically possible (see Equations 18, 32). In other words, they are not consistent with the extant laws of nature (note that this is a different claim from logical possibility). Recall that both naturalistic dualism and illusionism can be understood to be opposing responses to the same basic paradox of phenomenal judgment: phenomenal judgments seem to not require phenomenal consciousness. This opens the possibility for zombies. Chalmers (2017, 1996) takes the threat of zombies to force us to accept that phenomenal properties are an additional fundamental component of reality. Illusionists instead take *us* to be the zombies (Frankish, 2017; Graziano et al., 2020). But we have demonstrated that zombies are not nomologically possible in a relativistic universe. The illusionist can, of course, point out that we have assumed that Alice is phenomenally conscious, which they do not grant. While this is an assumption (and not a controversial one for any but the illusionist), what we point out is that the illusionist contention that there are zombies, and that *we* are the zombies, is premised on the paradox of phenomenal judgment. But we have eliminated that paradox by eliminating the nomological possibility of zombies. Therefore, we have eliminated the illusionist motivation to deny that Alice is phenomenally conscious. While we can't prove that Alice is indeed phenomenally conscious, the illusionist no longer has a reason (viz., the paradox of phenomenal judgment) to doubt it.

As a result of the relativity principle, there is a formal equivalence between functional consciousness (creating phenomenal judgments) and phenomenal consciousness (qualia and eidetic structures; see Equation 35). Ultimately, we are returning to Nagel's (1974) definition of consciousness: that the creature is conscious if there is something that is like to be this creature. If an observer has phenomenal judgments, then the observer measures that there is something that it is like to be that observer, and this observer is conscious. The essence of the relativistic principle is that there is nothing over and above the observer. Hence, in contrast to illusionism and naturalistic dualism, there cannot be any third-person or God's-eye perspective telling us that some observer is delusional and doesn't *really* have consciousness.

Moreover, phenomenal properties are no longer absolute determinations, but depend on the observer's cognitive frame of reference, just as in the case of constant velocity and the question of who's stationary and who's moving. In a non-relativistic universe, an object O evinces either property P or $\neg P$. Properties are absolute determinations. But in a relativistic universe, an object O evinces either property P or $\neg P$ *with respect to some observer X*. In the first-person cognitive frame of reference, quale Q (for example) is observable, while any third-person cognitive frame of reference observes $\neg Q$, i.e., that there is no quale.

Cognitive frames of reference are defined by the dynamics of the cognitive systems involved (according to the three-tier information processing model, eqs. 2–10). The first-person frame of reference takes into account also the position of the dynamics (eqs. 49–50). The first-person frame of reference of Alice, for example, is the position from within Alice's cognitive system, which is a physical position in space and time where the dynamics of the system take place. This is the only situation that satisfies the condition for the first-person frame of reference (eq. 42). For that reason, the transformation function $\Lambda^{\nu\mu}$ (that transforms between cognitive frames of reference and between first and third frames of reference) takes into account not only whether the input is recognizable by the cognitive system as one of its own representations, but also if the input is in the position of the cognitive system (eq. 46). The third person frames of reference are the positions of other cognitive systems, like Bob, that measure Alice's cognitive system. While Alice may observe herself feeling happiness as she's looking at a rose, Bob will only measure patterns of neural activity. Recall the case of constant velocity, wherein Alice claims to be at rest while Bob is moving, while Bob claims that Alice is the one moving and that he is stationary. From Alice's perspective, she has qualia and Bob only has patterns of neural activation, while from Bob's perspective Alice just has patterns of neural activation while he has qualia. In other words, just as in the case of constant velocity, Bob and Alice both measure that they are the stationary ones, and hence will not agree on who is stationary and who is moving, there is likewise a relativistic equivalence between all cognitive systems that have phenomenal consciousness or functional consciousness. Both Alice and Bob will measure that they are the ones who have phenomenal consciousness while the other has only neural patternsand hence will not agree on who has phenomenal consciousness.

Alice and Bob will continue to argue over who is right. As a result, Alice might be an illusionist and claim not only that Bob is delusional about having qualia, but moreover that qualia don't exist at all. And, as a response, Bob might claim that qualia are uniquely private and non-physical phenomena. But, just as with the constant velocity case, their conclusions are wrong because they don't grasp the relativistic principle. The illusionist mistake is to claim that the third-person perspective is the only legitimate perspective. When Alice observes Bob's neural firing, he supposedly should infer that all his "qualia" really are just neural firing. Yet the first-person perspective is also a physical arrangement of a cognitive system. It is the cognitive system from its own observer perspective. Both frames of reference are equivalent, there is no observer position that is privileged, and it is impossible to privilege one observer cognitive position over the other. Alice's perspective is not some absolute, "correct" perspective that sets the standard for measurement. Bob doesn't have phenomenal consciousness relative to Alice's perspective but *does* have phenomenal consciousness relative to Bob's perspective. In other words, Bob evinces $\neg Q$ with respect to Alice, and Bob evinces Q with respect to Bob. This is not a logical contradiction, but a statement of the relativity of properties.

On the other hand, qualia and eidetic structures are not private and hence phenomenal consciousness is neither some mysterious force beyond the realm of science nor an irreducible element of reality. Rather, they *appear* to be private because in order to measure them, one needs to be in the appropriate frame of reference, viz., that of the cognitive system in question (see Equations 42, 45). It is ultimately a question of causal power. Only from this frame of reference is there causal power for the representations in the dynamics of the system. Only from the frame of reference of the cognitive system are these neural patterns recognized as representations (eq. 45). These representations are the input and output of the cognitive system (eq. 2-10); they are the ones that cause the dynamics in this cognitive frame of reference. From outside that observer reference frame, as in the position the neuroscientist takes as a third-person observer, the same exact phenomena appear as neural computations. This is because the third-person perspective is constitutively outside of the dynamics of representations of the cognitive system in question, and hence these representations do not have any causal power over the neuroscientist (Equations 36–38). According to the equivalence principle, when Alice observes herself to be happy, it is because her cognitive system can recognize and use the appropriate phenomenal judgments, and these judgments are equivalent to phenomenal consciousness (Equation 35), because there is always an equivalent phenomenal judgment representation for every phenomenal property, and we cannot distinguish between them. These representations and their relations cause the cognitive dynamical system to react with new representations and instantiate new relationships between them. As a result, we get a dynamical system that uses *very specific* representations as variables and as outputs. Any other cognitive system, like that of Bob, uses *different* representations and thus cannot use the representations of Alice directly in its own dynamics (Equations 36–38). The only possibility left for Bob is to process Alice's representations through his sensory system and build his own representations. Consequently, we get a sharp difference between the self-measurements of the cognitive system (first-person perspective) and measurements from the outside the cognitive system (third-person perspective), which are mediated solely through the three-tier hierarchy—from the sensation subsystem to the perception subsystem and on to the functional consciousness subsystem (Equations 39, 40). Ultimately, the reason for the sharp difference between first-person and third-person frames of reference is because the sensation subsystem can observe only the physical properties (y_j) of inputs from the outside of the cognitive system, while the functional consciousness subsystem can only observe representations from within the cognitive system (Equations 41–43). Even if the two cognitive systems at hand are in the same cognitive frame of reference (which mean that they have the same dynamics), because they have different spatial positions where their dynamics take place, they will not be in the same first-person frame of reference. They will still need to use their sensation subsystem to measure the other system (hence satisfying the condition of the third-person frame of reference, eq. 43).

Phenomenal properties are not truly private. They seem private because it is non-trivial to do the transformation to the appropriate cognitive frame of reference. Nevertheless, this kind of transformation is nomologically possible. We showed that, as a result of the relativity principle, there is a transformation between the qualia of all cognitive frames of reference (Equations 44, 46, 51). Using this transformation, we obtained an equation that agrees with the relativistic principle (Equation 44). It describes the laws of physics with the same form in all admissible cognitive frames of reference. In other words, this form will stay the same regardless of which cognitive frame we choose (it doesn't matter for the equation what the specific μ and ν cognitive frames are). The equation enables us to move from the phenomenal consciousness of one cognitive frame of reference to the phenomenal consciousness of another frame and from first-person frame of reference to third person frame of reference (and vice versa). According to the transformation, each observer can change cognitive frames of reference to any other frame by changing the dynamics of its cognitive system. The transformation equation is built in such a way that ensures that after we applied the transformation and moved from one frame to the other, we will get the correct values of the measurements in the new cognitive frame of reference. For example, the equation ensures that the relations between first- and third-person frames of reference are always satisfied. The observer will always measure qualia and eidetic structures from within its own frame, and brain patterns (or any other physical patterns that govern the cognitive system) for all other frames. The only way to enter the frame of reference of the cognitive system is to have the dynamics of representations of that system, because they only have the right kind of causal power from within that cognitive system (they are the right "fuel" of the cognitive system – the inputs and outputs of the system). Hence, third-person studies will get different measurements than those taken from within the system itself, unless a proper transformation were to change the frame of reference of the third-person cognitive system to have the exact dynamics of the cognitive system in question (see Equations 36–38 for third-person frame, in contrast to Equations 42, 45 for the first-person frame). After such a transformation, there would be an equivalence between the two systems and they would observe the same conscious experience. Obviously, no such transformation is currently technologically feasible, but it is nomologically possible. Because of this transformation that preserves the form of the measurements, phenomenal structures (first-person perspective) and phenomenal judgement-representations (third-person perspective) are equivalent and isomorphic to each other. They are not equal, but they have the same preserved mathematical form such that we can map between them. All that separates them are different kinds of measurements (causing different kinds of properties).

The Hard Problem Dissolves

The relativistic theory of consciousness dissolves the hard problem. If there is no irreducibly private property of phenomenal features, there is no need for adding new, exotic elements to reality, nor is there a need to explain phenomenal features away as illusory. There is also no explanatory gap, because there is no need to explain how physical patterns in the brain create private, irreducible properties. Because consciousness is a relativistic phenomenon, physical patterns (e.g., of neural representations) and phenomenal properties (e.g., qualia) are two sides of the same coin. Both are valid physical descriptions of the same phenomenon from different frames of reference. Although phenomenal properties supervene on physical patterns (the representations that we called 'phenomenal judgements'), they are not created by the physical patterns. Instead, phenomenal properties are the result of a special measurement of these physical representations by the observer. For this observer, these representations are the ones that cause its own dynamics. Notice that according to the relativistic theory of consciousness, the opposite is also true and physical patterns (phenomenal judgements) supervene on phenomenal properties. They are not created by phenomenal properties; instead, physical representations are the result of a measurement of these phenomenal properties by a different observer. For this cognitive frame of reference, these phenomenal properties have no causal power over its dynamics. Phenomenal and physical aspects supervene on each other. There is a subtle identity between them (i.e., they are equivalent). They are just different perspectives of the same phenomenon from different cognitive frames of reference. As a result, the relativistic theory of consciousness is a physicalist theory, but not a reductive theory, because there is no reduction of phenomenal properties to brain patterns. For phenomenal properties we need a cognitive system with the right kind of representations and the right kind of measurement.

In the introduction, we raised a concern about whether any physicalist theory can solve the hard problem of consciousness. How can there be an identity between public properties like structure and dynamics and between private qualia? The answer is that we didn't include relativity as a part of physicalism. The relativistic theory of consciousness shows that phenomenal states are not private and gives an explanation of why they are different from other physical states. Different kinds of measurements give rise to different properties. While our sensation subsystem can only directly measure a physical substrate via measurement devices like the retina, the perceptual and cognitive subsystems directly measure only the roles and relations of representations within their subsystems and cannot directly measure the physical substrate serving as the referent of their representations (that's why in the equations that describe the dynamics of the perceptual and cognitive subsystems, the inputs are representations and not the physical substrate of the representations, while in the sensation system the inputs are the physical substrates themselves like light or sound waves). From these different kinds of measurements different kinds of properties arise.

Chalmers (1996, 2017) parses the zombie argument in terms of the logical possibility of zombies. He would not deny that zombies are inconsistent with the extant laws of nature, because he speaks only about logical possibility in general. However, the relativistic theory of consciousness carries ontological constraints and conditions about the existence of consciousness in every possible world. For example, the difference between first- and third-person frames of reference arise because of the difference between the kind of measurements that can be taken within each frame of reference. There could be a possible world, not governed by the known laws of nature, wherein the relativistic principle is false. But then, there would be no different frames of reference, and the different kinds of measurements would yield the same result. In other words, there would be no first-person and third-person frames of reference. In such a possible world, we would be left with either the illusionist view, with no consciousness at all (i.e., everybody is a zombie with just third-person perspective), or we would be left with a Berkeleian idealist world where there are only phenomenal properties (with shared first-person perspective for everything). Consequently, according to our relativistic theory of consciousness, there cannot be a world that has both phenomenal consciousness *and* zombies. This conclusion undermines the zombie argument. The main goal of the zombie argument was to establish that phenomenal properties cannot be explained reductively in terms of physics. But the relativistic theory does just that, demonstrating that phenomenal consciousness and physics can be reconciled. There might be a logical possibility of zombies, but in such possible worlds, the physical mechanism that allows the transformation to phenomenal properties (i.e., the relativistic principle) must be absent and hence consciousness must be absent all together as well (because all there is left in such a world is the physical substance that constitutes the world). In sum, in non-relativistic worlds without different frames of reference, we could have *either* illusionism (only physical properties) *or* Berkeleian idealism (only phenomenal properties). But in a relativistic world where different frames of reference measure different properties, *both* physical *and* phenomenal properties are possible for the same entity.

Phenomenal consciousness is not private—all we need to do to measure phenomenal properties is to change our perspective by moving to the appropriate frame of reference, i.e., the frame wherein the representations have causal power (i.e., the frame that measures the representations directly). As a result, this frame doesn't measure them merely as physical patterns, but as what they represent and stand for. For example, the representation of an apple has some causal power in the system, according to what has learned about apples. The representation will trigger, for example, memories, emotions, and motor processes. For the cognitive system, this is what it means to be an apple. Thus, the physical properties of the representation are not being directly measured, but rather only its role and relations with other representations in the system. Consequently, this representation will be measured by the system as an apple and not as a representation of an apple. That is to say, because there is nothing above and beyond the observer, and because the cognitive system (the observer) measures this representation according to what it does (the representation describes all the functions, properties and relationships of an apple in the system), this representation will be measured by the system as an apple. (That's why phenomenal consciousness is characterized by transparency, i.e., we don't perceive representations but directly perceive things). This is a direct measure of the representation itself. In the case of phenomenal judgments representations, these new properties are being measured as phenomenal properties. Specifically, when there is a cognitive frame of reference that creates functional consciousness (phenomenal judgments), it is sensible to assume that such a system will have a special subsystem specializing in such complex representations (the functional consciousness subsystem, *C*). It uses the complex representations of functional consciousness for its dynamics (i.e., $\vec{c}_{x_t}$ – phenomenal judgment representation, see eq. 5). As a result, these representations have causal power in this frame of reference. The relativity principle ensures us that because no measurement can distinguish between phenomenal properties and their corresponding phenomenal judgments, this cognitive frame will measure these phenomenal judgments as qualia and eidetic structures (Equation 35). In sum, according to the relativistic theory of consciousness, for phenomenal structures we need a cognitive system with the right kind of representations and the right kind of measurement. Phenomenal structures are the direct measurements of the complex representations that we called phenomenal judgments.

The relativistic theory of consciousness suggests a solution for the hard problem based in relativistic physics. There are still several open questions that need to be addressed in the future. For example, what are the necessary and sufficient conditions for a cognitive frame of reference to have phenomenal consciousness? And what empirical predictions does this theory yield? On the philosophical side, there are also questions about the affinities and differences between our view and dual aspect monisms such as reflexive monism (Velmans, 2009). These are issues that we will address looking forward.

The Formal Equivalence of Phenomenal and Neurocomputational Structures

Since consciousness is relativistic, there is a formal equivalence between first-person phenomenological structures and third-person neurocomputational structures. Philosophers have asked and puzzled over the question, "What is it like to be a bat?" (Nagel, 1974). Nagel's point is that physicalism is not sufficient for explaining consciousness. Even if we describe the neural processes involved in bat sonar, we will never get an intuitive or imagistic sense of what it is like from the first-person perspective. Yet years of philosophical arguments about the impossibility of a complete scientific understanding of consciousness has only stymied serious research, relegating consciousness to something that is supposedly not measurable and hence not amenable to scientific treatment (Holland, 2020).

By the relativistic theory of consciousness, the neurocomputational structures are equivalent to the phenomenal structures of what it is like to be a bat. This equivalence is not a statement of the reductive physicalism that Nagel attacks. We do not maintain that phenomenology can be reduced to neural computation. Rather, we maintain that phenomenology and neural computation are two different ways that the same phenomenon appears based on the cognitive perspective of the observer. That perspective is either from within the cognitive system (the first-person perspective) or outside of the cognitive system (the third-person perspective). This equivalence allows us to use neurocomputational structure to derive phenomenal structure. In other words, phenomenal consciousness can be investigated by studying neurocomputational dynamics. Certainly, the only way that the intuitive, imagistic aspects of "what it is like" could be perceived is by actually taking the cognitive frame of reference of the system in question. But qualia and eidetic structures can be measured indirectly through the neural phenomena they manifest from the third-person cognitive frame of reference. Phenomenal consciousness is not an immaterial and extra-scientific phenomenon, but one that's amenable to the scientific method (for an example of a similar approach, see Petitot, 2018).

Not only can neurocomputational structure help us discover phenomenological structure. Furthermore, phenomenology can also help us to discover neurocomputational structures themselves. We are referring to neurophenomenology (Varela, 1999; Gallagher, 2003; Varela and Thompson, 2003; Petitmengin, 2006; Petitot, 2018; Gallagher and Zahavi, 2020). Rigorous first-person philosophical inquiry into phenomenal structure can help guide cognitive neuroscience research into consciousness by tracing the lineaments of what scientists ought to be looking for. While descriptions of phenomenal consciousness are often thought of as wispy and vague, there are precise ways that phenomenological philosophy (and so-called "cognitive phenomenology") describes phenomenal consciousness that go beyond naïve, folk conceptions of experience.

For the illusionist, who understands phenomenal consciousness to be a cognitive illusion of phenomenal properties, "cognitive scientists should treat phenomenological reports as fictions" (Frankish, 2017, p. 27). In the relativistic perspective, however, phenomenological reports are not fictions. Of course, there are serious methodological issues that make data collection from the first-person perspective the purview of philosophy rather than science (Husserl, 1925/2012). But phenomenal reports are reports of what the system is like from within its own cognitive frame of reference. These reports can potentially help guide cognitive neuroscientific investigation. Since consciousness is relativistic, phenomenal structure is simply a different way of perceiving the same phenomenon that the cognitive neuroscientist is examining from a different observer perspective. Rather than denigrating the first-person perspective, we believe the relativistic framework invites us to take phenomenological philosophy seriously. We believe it poses potentially rich possibilities for interdisciplinary work on consciousness. No doubt, such interdisciplinary work would be fraught with difficulties, but the phenomenological philosopher looking at phenomenal consciousness and the cognitive neuroscientist looking at functional consciousness are both looking at the same thing from different observer positions.

Let's get a clearer idea on what philosophers can contribute to the empirical study of consciousness. Typically, first-person experience is spoken of in terms of qualia, which often picks out the content of phenomenal consciousness: what it is like to see red (Jackson, 1982), or what it is like to be a bat and have sonar (Nagel, 1974). But the tradition of phenomenological philosophy, initiated by Edmund Husserl and continued by Martin Heidegger and Maurice Merleau-Ponty, has long studied first-person experience in terms of eidetic structures. An eidetic structure refers to the "form" or "essence" (*eidos*) of a given phenomenon in first-person experience, that is, its invariant structure across time. Phenomenologists are interested not in the qualia of greenness, but in the structural invariants that are expressed in all phenomenal consciousness, such as intentionality or the temporal structure of cognition and consciousness. We will close this section by describing the eidetic structure of what phenomenologists call time-consciousness (Husserl, 1917/1991; Neemeh and Gallagher, 2020). We will hypothesize a way that this eidetic structure may be realized as a neurocomputational structure. This should be understood as an invitation for the kind of interdisciplinary exploration between philosophers and cognitive neuroscientists that we have advocated for.

The subjective experience of time has a formal eidetic structure, what Husserl (1917/1991) calls time-consciousness. While we usually imagine time to be a linear arrow, phenomenologists contend that time-consciousness has a far more complicated, looping structure (Neemeh and Gallagher, 2020). Take the perception of a melody. How can you perceive a melody if, at any given point in time, you only hear a single note being played? Don't you simply perceive individual notes? But a series of individual notes is not the same as hearing a melody. The eidetic structure of melody perception involves what phenomenologists call protention and retention. First of all, the immediate present is never a simple, non-extended slice of time. The present of experience is a "specious present" (James, 1890/1983) that is temporally extended. When one note is perceived in the immediate present, it recedes into the immediate past. Yet it is still retained, in an attenuated form, in consciousness (called "retention"). It is retained not as a sensuously fulfilled object, but as a trailing shadow (Husserl, 1917/1991). Only the immediate now is fully sensuous. But if past notes were not retained in this way in consciousness, we would not have a conscious perception of a melody. We would only hear individual notes. At the same time, melody perception involves an anticipation (or protention) of the imminent note. This is how we can perceive a wrong note: it doesn't match up with our anticipation, based on the previously given aspects of the melody. The same principle holds when we hear a sentence.

In the relativistic view, first-person experience is no longer ineluctably private. It is simply one specific cognitive frame of reference, and when we switch perspectives, by the appropriate transformation (eq. 44, 46, 51), to that of cognitive neuroscience (third-person perspective), we ought to (ideally) detect isomorphic structures that are not qualitative but rather quantitative (assuming our phenomenological analysis is correct, of course).

Given this formal equivalence (Equation 44), there is a neurocomputational structure that is equivalent to the eidetic structure of time-consciousness, or perhaps there are multiple such neurocomputational structures for time-consciousness in different sensory modalities. Cognitive neuroscientific investigation can use this eidetic structure as a guide to inquiry. For example, Varela (1999) suggests an account based in phase locking, and Neemeh and Gallagher (2020) suggest a Bayesian approach.

Philosophers have long studied phenomenal consciousness and may have practiced insight into what structures to look out for in functional consciousness. This is not to say that philosophy is, or has ever pretended to be, a science. We advocate for an interdisciplinary investigation of consciousness that takes eidetic structures as the seeds of empirical hypotheses into neural function (qualia, on the other hand, would likely not be very productive of hypotheses. After all, how much is there to say about redness?). This can open the cognitive neuroscience of consciousness up to vistas of richer views of consciousness.

CONCLUSION

In sum, we propose a novel, relativistic theory of consciousness, one that accounts for both the functional and phenomenal features of consciousness, bridging the explanatory gap. Through conceptual arguments and mathematical formalizations, we propose that there is no need to expand the basic inventory of nature (as dualists like Chalmers, 2017 argue), nor is there a need to explain away phenomenal features (as illusionists argue). Phenomenal features are not truly private, since the principle of relativity allows us to perform a transformation from one cognitive frame of reference to another. We provided a mathematic transformation between two idealized cognitive systems taken from different cognitive frames of reference, showing their relativistic equivalence. The privacy of phenomenal features is only an illusion, based on our biological limitations and the technological limitations of current science—basically, we can't yet actually perform such a transformation. But our formalization is a proof of concept, showing that it is theoretically feasible. Since phenomenal features are not private, both the presence of zombies and the paradox of phenomenal judgment fall away. The dualist infers from these that phenomenal consciousness is a non-material extra force or property of nature, while the illusionist infers that phenomenal consciousness is merely an illusion created by phenomenal judgments. But once the privacy issue, zombies, and this paradox are neutralized, there is no longer any strong motivation for the dualist and illusionist positions. Phenomenal consciousness is neither private nor delusional, just relativistic.

Not only does the relativistic theory of consciousness legitimize the study of phenomenal features in science, but it furthermore opens up many new questions and possibilities for research. We noted that philosophers studying phenomenal consciousness could play a legitimate role in the science of consciousness, such as by theoretical contributions to experiments seeking the neural basis of phenomenal and eidetic structures.

DATA AVAILABILITY STATEMENT

The original contributions presented in the study are included in the article/**Supplementary Material**, further inquiries can be directed to the corresponding author.

AUTHOR CONTRIBUTIONS

Both authors listed have made a substantial, direct, and intellectual contribution to the work, and approved it for publication.

ACKNOWLEDGMENTS

We would like to acknowledge the Kavli Institute for Cognitive Neuroscience, which generously hosted and led us to develop the ideas in this article, as well as the suggestions of reviewers. NL would like to thank Ashok Vaish for his continuous and generous support. NL would also like to thank Benjamin Pageler, Eva Jablonka, Simona Ginsburg, Avshalom Elitzur, and David Steinhart.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fpsyg.2021.704270/full#supplementary-material

REFERENCES

Alter, T. (2016). The structure and dynamics argument against materialism. *Noûs* 50, 794–815. doi: 10.1111/nous.12134

Baars, B. J. (1988). *A Cognitive Theory of Consciousness*. Cambridge: Cambridge University Press.

Block, N. (2011). Perceptual consciousness overflows cognitive access. *Trends Cogn. Sci.* 15, 567–575. doi: 10.1016/j.tics.2011.11.001

Brown, R. (2010). Deprioritizing the a priori arguments against physicalism. *J. Conscious. Stud.* 17, 47–69.

Block, N. (1995). On a confusion about a function of consciousness. *Behav. Brain Sci.* 18, 227–247. doi: 10.1017/S0140525X00038188

Carroll, S. (2021). Consciousness and the laws of physics. *J. Conscious. Stud.* 28, 16–31.

Chalmers, D. (2017). "Naturalistic dualism," in *The Blackwell Companion to Consciousness*, eds S. Schneider and M. Velmans (Chichester: John Wiley & Sons), 363–373. doi: 10.1002/9781119132363.ch26

Chalmers, D. J. (1995). Facing up to the problem of consciousness. *J. Conscious. Stud.* 2, 200–219.

Chalmers, D. J. (1996). *The Conscious Mind: In search of a Fundamental Theory*. Oxford: Oxford university press.

Chalmers, D. J. (2003). "Consciousness and its place in nature," in *Blackwell Guide to the Philosophy of Mind*, eds S. Stich and T. Warfield (Oxford: Blackwell), 102–142. doi: 10.1002/9780470998762.ch5

Dehaene, S. (2014). *Consciousness and the Brain: Deciphering How the Brain Codes Our Thoughts*. New York: Penguin.

Dennett, D. C. (1991). *Consciousness Explained*. New York: Penguin.

Dennett, D. C. (1995). The unimagined preposterousness of zombies. *J. Conscious. Stud.* 2, 322–326.

Doerig, A., Schurger, A., Hess, K., and Herzog, M. H. (2019). The unfolding argument: why IIT and other causal structure theories cannot explain consciousness. *Conscious. Cogn.* 72, 49–59. doi: 10.1016/j.concog.2019.04.002

Dretske, F. (1978). *The Role of the Percept in Visual Cognition*. Minneapolis: University of Minnesota Press.

Dretske, F. (2003). "Sensation and perception (1981)," in *Essays on Nonconceptual Content*, ed. Y. H. Gunther (Cambridge: The MIT Press).

Einstein, A. (1905). On the electrodynamics of moving bodies. *Ann. Der Phys.* 17, 891–921. doi: 10.1002/andp.19053221004

Einstein, A. (1907). *Über das Relativitätsprinzip und die aus demselben gezogenen Folgerungen*. Stuttgart: S. Hirzel.

Einstein, A. (1911). On the influence of gravitation on the propagation of light (M. D. Godfrey, Trans.). *Ann. Phys.* 35, 898–908.

Einstein, A., Lorentz, H. A., Minkowski, H., Weyl, H., and Sommerfeld, A. (1923/1952). *The Principle of Relativity: A Collection of Original Memoirs on the Special and General Theory of Relativity*. Illinois: Dover. (Original work published 1923).

Fahrenfort, J. J., Scholte, H. S., and Lamme, V. A. (2007). Masking disrupts reentrant processing in human visual cortex. *J. Cogn. Neurosci.* 19, 1488–1497. doi: 10.1162/jocn.2007.19.9.1488

Forshaw, J., and Smith, G. (2014). *Dynamics and Relativity*. Hoboken: John Wiley & Sons.

Frankish, K. (2007). The anti-zombie argument. *Philos. Q.* 57, 650–666. doi: 10.1111/j.1467-9213.2007.510.x

Frankish, K. (ed.) (2017). *Illusionism As a Theory of Consciousness*. United Kingdom: Andrews UK.

Franklin, S., Madl, T., Strain, S., Faghihi, U., Dong, D., Kugele, S., et al. (2016). A LIDA cognitive model tutorial. *Biol. Inspired Cogn. Archit.* 16, 105–130. doi: 10.1016/j.bica.2016.04.003

Gallagher, S. (2003). Phenomenology and experimental design toward a phenomenologically enlightened experimental science. *J. Conscious. Stud.* 10, 85–99.

Gallagher, S., and Zahavi, D. (2020). *The Phenomenological Mind*, 3rd Edn. Oxfordshire: Routledge.

Graziano, M. S. (2019). *Rethinking Consciousness: A Scientific Theory of Subjective Experience*. New York: WW Norton & Company.

Graziano, M. S., Guterstam, A., Bio, B. J., and Wilterson, A. I. (2020). Toward a standard model of consciousness: reconciling the attention schema, global workspace, higher-order thought, and illusionist theories. *Cogn. Neuropsychol.* 37, 155–172. doi: 10.1080/02643294.2019.1670630

Holland, O. (2020). Forget the Bat. *J. Artif. Intell. Conscious.* 7, 83–93. doi: 10.1142/S2705078520500058

Husserl, E. (1917/1991). *On the Phenomenology of the Consciousness of Internal Time*. Netherlands: Kluwer.

Husserl, E. (1925/2012). *Phenomenological Psychology: Lectures, Summer Semester, 1925*. Berlin: Springer Science & Business Media.

Hutto, D. D. (2007). The narrative practice hypothesis: origins and applications of folk psychology. *R. Inst. Philos. Suppl.* 60, 43–68. doi: 10.1017/cbo9780511627903.004

Jackson, F. (1982). Epiphenomenal qualia. *Philos. Q.* 32, 127–136. doi: 10.2307/2960077

James, W. (1890/1983). *The Principles of Psychology*. Cambridge: Harvard University Press.

Kanai, R., Chang, A., Yu, Y., Magrans, de Abril, I., Biehl, M., et al. (2019). Information generation as a functional basis of consciousness. *Neurosci. Conscious.* 2019:niz016. doi: 10.1093/nc/niz016

Kanizsa, G. (1985). Seeing and thinking. *Acta Psychol.* 59, 23–33. doi: 10.1016/0001-6918(85)90040-X

Lamme, V. A. (2006). Towards a true neural stance on consciousness. *Trends Cogn. Sci.* 10, 494–501. doi: 10.1016/j.tics.2006.09.001

Levine, J. (1983). Materialism and qualia: the explanatory gap. *Pac. Philos. Q.* 64, 354–361. doi: 10.1111/j.1468-0114.1983.tb00207.x

Mashour, G. A., Roelfsema, P., Changeux, J.-P., and Dehaene, S. (2020). Conscious processing and the global neuronal workspace hypothesis. *Neuron* 105, 776–798. doi: 10.1016/j.neuron.2020.01.026

Mindt, G. (2017). The problem with the'information'in integrated information theory. *J. Conscious. Stud.* 24, 130–154.

Møller, C. (1952). *The Theory of Relativity*, 2nd Edn. Oxford: Oxford University Press.

Nagel, T. (1974). What is it like to be a bat?. *Philos. Rev.* 83, 435–450. doi: 10.2307/2183914

Nagel, T. (2012). Conceiving the impossible and the mind-body problem. *Rev. Română Filosofie Anal.* 6, 5–21.

Neemeh, Z. A., and Gallagher, S. (2020). "The phenomenology and predictive processing of time in depression," in *The Philosophy and Science of Predictive Processing*, eds D. Mendonça, M. Curado, and S. Gouveia (London: Bloomsbury), 187–207.

Oizumi, M., Albantakis, L., and Tononi, G. (2014). From the phenomenology to the mechanisms of consciousness: integrated information theory 3.0. *PLoS Comput. Biol.* 10:e1003588. doi: 10.1371/journal.pcbi.1003588

Pageler, B. T. (2011). The Explanatory Value of the Cognitive Unconscious. Ph.D. dissertation.

Petitmengin, C. (2006). Describing one's subjective experience in the second person: an interview method for the science of consciousness. *Phenomenol. Cogn. Sci.* 5, 229–269. doi: 10.1007/s11097-006-9022-2

Petitot, J. (2018). *Elements of Neurogeometry*. Berlin: Springer.

Pound, R. V., and Rebka, G. A. Jr. (1960). Apparent weight of photons. *Phys. Rev. Lett.* 4:337. doi: 10.1103/PhysRevLett.4.337

Pylyshyn, Z. W. (2003). *Seeing and Visualizing: It's Not What You Think*. Cambridge: MIT press.

Solms, M. (2021). *The Hidden Spring: A Journey to the Source of Consciousness*. New York, NY: WW Norton & Company.

Tononi, G. (2008). Consciousness as integrated information: a provisional manifesto. *Biol. Bull.* 215, 216–242. doi: 10.2307/25470707

Tononi, G., Boly, M., Massimini, M., and Koch, C. (2016). Integrated information theory: from consciousness to its physical substrate. *Nat. Rev. Neurosci.* 17, 450–461. doi: 10.1038/nrn.2016.44

van Gulick, R. (2014). "Consciousness," in *Stanford Encyclopedia of Philosophy*, ed. E. N. Zalta (Berlin: Spring).

Varela, F., Lachaux, J.-P., Rodriguez, E., and Martinerie, J. (2001). The brainweb: phase synchronization and large-scale integration. *Nat. Rev. Neurosci.* 2, 229–239. doi: 10.1038/35067550

Varela, F., and Thompson, E. (2003). "Neural synchrony and the unity of mind: A neurophenomenological perspective," in *The Unity of Consciousness*, ed. A. Cleeremans (Oxford: Oxford University Press).

Varela, F. J. (1999). "The specious present: A neurophenomenology of time consciousness," in *Naturalizing Phenomenology*, eds J. Petitot, F. J. Varela, B. Pachoud, and J.-M. Roy (California: Stanford University Press).

Velmans, M. (2009). *Understanding Consciousness*. Routledge.

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

frontiers | Frontiers in Psychology

ORIGINAL RESEARCH
published: 18 May 2022
doi: 10.3389/fpsyg.2022.811712

Check for updates

What Individuals Experience During Visuo-Spatial Working Memory Task Performance: An Exploratory Phenomenological Study

Aleš Oblak[1]*, Anka Slana Ozimič[2], Grega Repovš[2] and Urban Kordeš[3,4]

[1] Laboratory for Cognitive Neuroscience and Psychopathology, University Psychiatric Hospital Ljubljana, Ljubljana, Slovenia, [2] Mind & Brain Lab, Department of Psychology, Faculty of Arts, University of Ljubljana, Ljubljana, Slovenia, [3] Center for Cognitive Science, Faculty of Education, University of Ljubljana, Ljubljana, Slovenia, [4] Observatory: Laboratory for Empirical Phenomenology, Faculty of Education, University of Ljubljana, Ljubljana, Slovenia

OPEN ACCESS

Edited by:
Narayanan Srinivasan,
Indian Institute of Technology Kanpur, India

Reviewed by:
Terje Sparby,
Rudolf Steiner University College, Norway
Dario Marin,
Sapienza University of Rome, Italy

***Correspondence:**
Aleš Oblak
oblak.ales.93@gmail.com

Specialty section:
This article was submitted to Consciousness Research, a section of the journal Frontiers in Psychology

Received: 09 November 2021
Accepted: 30 March 2022
Published: 18 May 2022

Citation:
Oblak A, Slana Ozimič A, Repovš G and Kordeš U (2022) What Individuals Experience During Visuo-Spatial Working Memory Task Performance: An Exploratory Phenomenological Study. Front. Psychol. 13:811712. doi: 10.3389/fpsyg.2022.811712

In experimental cognitive psychology, objects of inquiry are typically operationalized with psychological tasks. When interpreting results from such tasks, we focus primarily on behavioral measures such as reaction times and accuracy rather than experiences – i.e., phenomenology – associated with the task, and posit that the tasks elicit the desired cognitive phenomenon. Evaluating whether the tasks indeed elicit the desired phenomenon can be facilitated by understanding the experience during task performance. In this paper we explore the breadth of experiences that are elicited by and accompany task performance using in-depth phenomenological and qualitative methodology to gather subjective reports during the performance of a visuo-spatial change detection task. Thirty-one participants (18 females) were asked to remember either colors, orientations or positions of the presented stimuli and recall them after a short delay. Qualitative reports revealed rich experiential landscapes associated with the task-performance, suggesting a distinction between two broad classes of experience: phenomena at the front of consciousness and background feelings. The former includes cognitive strategies and aspects of metacognition, whereas the latter include more difficult-to-detect aspects of experience that comprise the overall sense of experience (e.g., bodily feelings, emotional atmosphere, mood). We focus primarily on the background feelings, since strategies of task-performance to a large extent map onto previously identified cognitive processes and discuss the methodological implications of our findings.

Keywords: visuo-spatial working memory, empirical phenomenology, psychological task, constructivist grounded theory, background feelings

INTRODUCTION

In experimental cognitive psychology, the objects of study are usually operationalized with psychological tasks. Many readers have probably participated in studies with such tasks, sitting in front of a computer screen with the typical light gray background, and waiting for stimuli to appear. In well-established domains of research, these tasks are standardized so that they can be easily modified to answer specific research questions. When interpreting the results of psychological tasks, we focus primarily on behavioral measures such as reaction times and performance accuracy rather than phenomenology. We assume that the psychological tasks elicit the desired cognitive phenomenon (cf. Morrison et al., 2019) and that we can infer underlying cognitive mechanisms based on behavioral performance.

For example, when participants perform a working memory task—where working memory is considered one of the most important cognitive functions, encompassing many complex cognitive operations (Baddeley and Hitch, 1974; Repovš and Baddeley, 2006; Baddeley, 2010)—it is assumed that the phenomenon that the task is intended to elicit (i.e., working memory) is the primary content of participants' conscious awareness. Based on this assumption, behavioral data are used to make inferences about the cognitive mechanisms of working memory (for an overview of this type of reasoning, see Caramazza, 1986; De Hollander et al., 2016; Coltheart, 2017).

If we are interested in measuring the target phenomena, we need to examine the validity of the task: understood here in the broadest sense as whether the task elicits the phenomenon in question (Stone, 2019). Assessing validity and interpreting behavioral data can be facilitated by understanding the experience while performing the psychological task (cf. Jaspers, 1997). In cognitive science, phenomenology has traditionally been used to study phenomena in naturalistic settings (e.g., Hurlburt, 1990; Gallagher et al., 2015; Oblak, 2020). When subjective reports are collected in laboratory settings, it is typically done to ensure that the task is working as intended. For example, Nelson et al. (2003) used a verbal change-detection task in which they manipulated the frequency of stimuli to elicit a familiarity-related conflict (e.g., a negative probe in the current trial was positive in the preceding trial). At the end of the experiment, a debriefing interview was used to determine whether participants were aware of this manipulation. More recently, however, in-depth phenomenological interviews have also been used with experimental tasks to assess whether the task elicits the desired phenomenology or to determine what phenomenology is associated with the task in the first place (for a theoretical overview, see Weger and Wagemann, 2015; Wagemann, 2021; for empirical examples, see Valenzuela-Moguillansky et al., 2013; Hurlburt et al., 2016).

Studies that combine phenomenological methods and psychological tasks focus primarily on task-related phenomenology. One aspect of task-related phenomenology that has been explored most are cognitive strategies. Seghier and Price (2018) propose that the same psychological tasks can be solved using different cognitive strategies (for empirical examples, see Coltheart et al., 1993; Tsukiura et al., 2005; Cummine et al., 2013; Braga et al., 2017). In spatial navigation tasks, to remember space, participants may rely on unitary encoding strategies (i.e., a single object is remembered based on cardinal directions), binary encoding strategies (i.e., objects are remembered based on the spatial relationships between them), or a combination of both (König et al., 2017). Questionnaire surveys revealed four cognitive styles that determine the strategy used to store stimuli in a memory retrieval task: verbal, visual, spatial, and image-based strategies (Miller et al., 2012); and two strategies used in performing a spatial working memory task: an auditory and a visuospatial strategy (Sanfratello et al., 2014). Furthermore, eye-tracker data distinguish between categorical and detail-based strategies when performing a spatial working memory task (Starc et al., 2017). In memory tasks, subtle changes in strategy, such as focusing on maintenance or retrieval, are associated with behavioral differences in performance (Speer et al., 2003). The use of different strategies in the same psychological task may imply that the task does not elicit the desired phenomenon and that the validity of the task may be in question (for discussions of the relationship between phenomenology, behavior, and neural dynamics, see Mcintyre, 1999; Mogensen and Overgaard, 2018).

To overcome this problem in the working memory domain, researchers use a variety of approaches, including the use of control tasks. For example, studies may be specifically interested in visual working memory and therefore want to control for the specific type of representation that participants are encoding. The possibility that participants encode stimuli in auditory form (i.e., by naming them in inner speech) rather than visual form is commonly controlled using so-called distractor tasks (e.g., articulatory suppression), in which participants must—in parallel with visual working memory task performance—continuously repeat a specific verbal phrase that prevents them from encoding and maintaining the target stimuli in auditory form (Barrouillet et al., 2007).

In addition to the various strategies participants experience when performing psychological tasks, psychological tasks are typically accompanied by a variety of confounding phenomenology, such as boredom (D'Angiulli and Smith LeBeau, 2002), anxiety (Ikeda et al., 1996), mind-wandering (Morrison et al., 2019), and task-related emotions (Laybourn et al., 2022). While working memory has been studied from a phenomenological perspective (Buchsbaum, 2013), to the best of our knowledge, novel approaches in first-person research (e.g., descriptive experience sampling, Hurlburt, 2011; micro-phenomenological interview, Petitmengin, 2006; for reviews of these methodological approaches, see Froese et al., 2011a,b; Valenzuela-Moguillansky and Demšar, 2022) have not been applied to this phenomenon. The full range of experiences that can be elicited by working memory tasks is therefore unknown.

The aim of this paper is to contribute to the understanding of working memory task performance by exploring the range of experiences evoked by and accompanying task performance. We consider such mapping of the space of experiences as a first step that can then guide further, more detailed and focused investigations of how the identified experiences may affect task performance and its neural correlates. By using a qualitative methodology to explore experiences during task performance, we aim to address four challenges: (i) the unreliability of current closed-form approaches (e.g., questionnaires, semi-structured debriefing interviews); (ii) limiting the exploration of confounding phenomena to a limited an incomplete set of *a priori* categories; (iii) incomplete interpretations of quantitative inferences in the absence of supporting qualitative findings; (iv) understanding the source of noise in the data. We discuss these aspects in more detail in the following paragraphs.

First, although there are several studies that attempt to understand differences in the strategies used in psychological tasks in general and working memory tasks in particular (see above), they usually rely on indirect measures. The existence of different strategies is inferred from behavioral or physiological data (e.g., Starc et al., 2017). Moreover, it is assumed that cognitive styles and strategies are sufficiently understood, so closed-form instruments (e.g., questionnaires) are often used to collect data on them. In recent decades, phenomenology has entered cognitive science (Varela et al., 1991; Flannagan, 1992; Thompson, 2007). Notably, an in-depth and systematic look at how individuals experience the world has been associated with a wide-range reexamination of what are the objects of inquiry in the sciences of the mind (for how phenomenology reexamined diagnostic criteria in the RDoC framework in psychopathology, for example, see Cuthbert and Insel, 2010). Such reexaminations have proven to be especially problematic for the use of closed-form questionnaires.

The second reason for choosing a qualitative methodology is that the assumption that phenomena accompanying psychological tasks are known may be invalid and needs to be empirically assessed. As noted earlier, studies using psychological tasks assume that participants experience only task-related cognitions when engaged in the task. However, as one study (Morrison et al., 2019) has shown, experiences during task performance can be quite different from what researchers intended—for a broader discussion of how the phenomenology of interaction with complex systems cannot be known *a priori* and requires further investigation, see theory-experience gap in Froese et al. (2012). Of particular interest are aspects of experience that are well known in the phenomenological tradition—e.g., existential feelings (Ratcliffe, 2008), background consciousness (Colombetti, 2011), and fringe awareness (James, 1890)—, but have not yet been explored in working memory research (although for a similar approach, see Laybourn et al., 2022). In-depth interviews can identify and provide insights into these and other aspects of experience that theory-driven questionnaires may fail to capture.

The third reason for choosing a qualitative methodology to investigate the experience during visuo-spatial working memory task performance is related to the ongoing replication crisis (Ioannidis, 2005). Recently, a new interpretation of the replication crisis has been proposed: the so-called generalizability crisis. Namely, the generalizability crisis stems in part from the fact that statistical claims in research papers about phenomena are incompatible with qualitative claims. It has been argued that commonly, there is no link between statistically testable experiments (presented in section "Materials and Methods") and the broader claims made about the phenomena of inquiry (presented in section "Introduction" and "Discussion"), and that this gap might be addressed with formal qualitative research (Yarkoni, 2022). In his seminal monograph on descriptive experience sampling, Hurlburt (2011, chapter 21) states that first-person reports should not serve merely as the initial, exploratory step in the study of a phenomenon. Rather, at each step of theory-construction and conducting experimental research, we should reevaluate whether we are still engaging with the target phenomenology or whether we have begun to engage with theoretical abstractions.

Finally, to understand and manage variability in behavioral and neural data, it is important to have a comprehensive overview of possible sources of "noise"— phenomena that accompany the cognitive process of interest (Seghier and Price, 2018). Our final aim, therefore, was to capture the breadth of experience during the performance of a working memory task in order to build a taxonomy of experience categories that can be used in future studies (cf. Lutz et al., 2002; Hurlburt et al., 2016; Fernyhough et al., 2018). Importantly, our aim was not to examine inter-subject variability along experimental dimensions, but rather to examine the breadth of different experiences that participants may have while performing the working memory task.

To date, only two studies have used modern approaches to phenomenal data collection to incorporate the study of experience into the analysis of task performance. First, it has been shown that during performance of simple visual tasks, different strategies accessible with subjective reports are associated with different electrophysiological signatures (Lutz et al., 2002). Second, elicited and spontaneously occurring inner speech are associated not only with different but also opposite patterns of neural activity recorded by fMRI (Hurlburt et al., 2016; Fernyhough et al., 2018). To investigate the experience during visuo-spatial working memory task performance, we used an in-depth phenomenological and qualitative methodology, in particular constructivist grounded theory (Charmaz, 2004). This approach aims to outline in detail the structure of a given phenomenon. Thus, it is interested in describing the widest possible range of different experiences associated with a phenomenon. To facilitate detailed descriptions, constructivist grounded theory gathers as diverse data as possible from as many different sources as possible. Thus, we collected in-depth qualitative data from a heterogeneous sample of participants with different ages, educational backgrounds, and experience in mind sciences in different visuo-spatial working memory tasks. The data collected suggested a clear distinction between two broad classes of experiences: phenomena at the forefront of consciousness and background feelings. The former includes strategies and aspects of metacognition related to task performance, whereas the latter include descriptions of the overall impression of the experience (e.g., bodily feelings, emotional atmosphere, mood). We present all experiential categories induced in this study. However, because task performance strategies largely correspond to task-related cognitive processes identified in previous research, we focus primarily on background feelings.

MATERIALS AND METHODS

A combination of a computerized behavioral task designed to investigate visuo-spatial working memory and an in-depth phenomenological interview was used to gather first-person data. Participants were asked to attend four interview sessions. During each session, participants were asked to solve multiple trials of the visuo-spatial working memory task. At a random moment during task-performance, participants were prompted to report on their experience. A phenomenological interview followed. The interview investigated both how participants' experience evolved through time during the trial immediately before the prompt, as well as what they experienced during each moment of the trial. Additionally, behavioral data on their task-performance (i.e., performance accuracy and reaction times) were gathered but are not reported in this paper and will be presented elsewhere.

Audio recordings of the interviews were analyzed according to the principles of constructivist grounded theory. Our main analytical instrument was coding (i.e., assigning general descriptive tags to sections of raw data). Three stages of coding took place. First, codes were induced solely from the data. Second, extant categories were compared to as-of-yet uncoded data, while novel categories were induced. Third, all the categories were fitted to the data. In line with principles of qualitative research, data acquisition and analysis were conducted in parallel, one informing the other. Then, relational coding was used to establish logical relationships between individual experiential categories constructed via coding. Finally, a codebook was constructed, outlining both individual experiential categories, and the relationships between them. In a subset of participants, the resulting codebook was validated in additional interviews.

Figure 1 provides a schematic outline of the research design. Each aspect of the research design is presented in detail in subsequent subsections.

Participants

Thirty-one participants (18 females) aged between 20 and 50 years (*M* = 27.0, *SD* = 6.21; years of education: *M* = 16.52, *SD* = 2.35) signed an informed consent to participate in four 60-min study sessions in which they performed a visuo-spatial working memory task. All participants had normal or corrected-to-normal eyesight. Participants did not self-report any neurological or psychiatric disorders. All but three participants were right-handed. The sample size was determined based on conceptual depth of the data (see section "Data Analysis"), and by the expectation that it would match or exceed the sample size in a typical working memory study (usually, between 20 and 30 participants, e.g., Bo and Seidler, 2009). 16 participants were students of cognitive science at the University of Ljubljana. Nine of them participated in the study as a part of their coursework. Since their participation was therefore not strictly speaking voluntary, they were given the option for their data to not be used in the analysis. None of the participants opted for having their data removed. To avoid the possibility of gathering data that are biased by theoretical understanding of either working memory or phenomenology, additional 15 participants who had no background in mind sciences were recruited. While quantitative studies ask for samples to be as homogenous as possible, the aims of grounded theory differ and a varied and heterogenous sample is desirable so as to account for as broad a range of dimensions associated with the phenomenon of inquiry as possible (cf. Charmaz, 2004). The nine participants who were involved in the study as a part of the course in first-person research received course credits. For non-student participants no reward was given in exchange.

Instruments

Visuo-Spatial Change Detection Task

The participants completed multiple trials of the visuo-spatial change detection task, in which they were asked to memorize orientations, colors or positions of the presented stimuli. The timing structure of each trial is presented in **Figure 2**. Briefly, in the remember color and remember orientation conditions, four target stimuli were presented simultaneously for 2.0 s. In the remember position, the target stimuli were presented sequentially, each for the duration of 0.75 s. Following a maintenance delay of 2.0 s, the probe stimuli were shown, and the participants had to indicate by a button press whether there was a change in the relevant property in any of the stimuli. Both the accuracy and reaction times of the responses were recorded. If the participant did not respond within 2.5 s, a null response was recorded, and the participant proceeded to the next trial.

The visuo-spatial change detection task was presented on an Acer Aspire 3 laptop (Intel Core i5 processor with 2.50 GHz and 3 MB RAM) running Windows 10 Pro operating system. The stimuli were presented on a 15.6-inch full LED screen with 1920 × 1080 resolution and 60 Hz refresh rate. The screen was set to maximum brightness (224 cd/m2) when the task was being performed. Participants sat approximately 75 cm away from the center of the screen.

A custom script prepared in PsychoPy (Peirce, 2007) was used to present the stimuli and collect responses. Stimuli in the orientation condition were black keys (see **Figure 2**), 35 px (0.55° visual angle) in length, with the main circle 19 px (0.3° visual angle) in diameter and the line 7 px (0.11° visual angle) wide. The orientation of each stimulus was randomly selected from eight possible principal directions pointing toward 0, 45, 90, 135, 180, 225, 270, and 315°. Stimuli in the color condition were colored circles, 35 px (0.55° visual angle) in diameter. The color of each stimulus was randomly selected from eight easily distinguishable color hues: red, dark blue, light blue, green, yellow, purple, and white. In the position condition, the stimuli were black circles, 20 px (0.32° visual angle) in diameter.

Stimuli in all conditions were presented on a gray background within an invisible square bounding box 520 px (8.14° visual angle) in width positioned in the center of the screen. The position of the orientation and color stimuli was fixed in the vertices of the invisible square. The exact position of the stimuli within the square varied randomly between trials, with the requirement that the minimal distance between the centers of each pair of objects be at least 4 times their bounding radius.

The Phenomenological Interview

The automated change-detection task was designed to allow pausing after any trial and allow for experiential sampling. Specifically, between the seventh and the fifteenth trial, the interviewer paused task execution after participants responded to the probe stimuli. The interviewer then asked the participants to stop performing the task and reflect on their

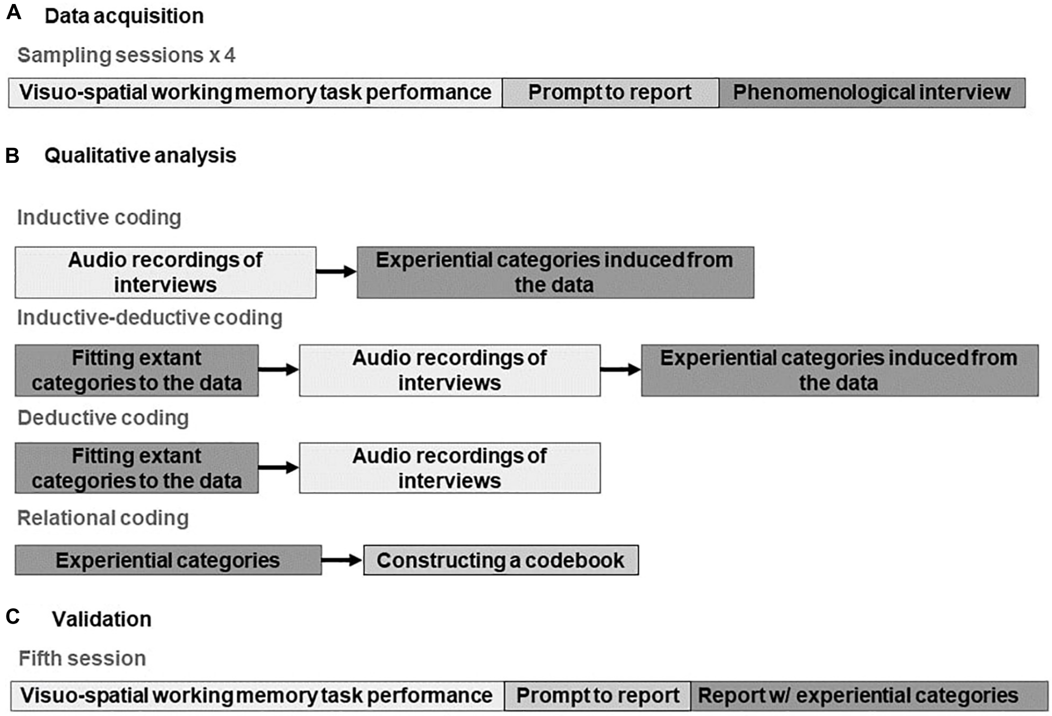


FIGURE 1 | (A–C) Overview of the research design.

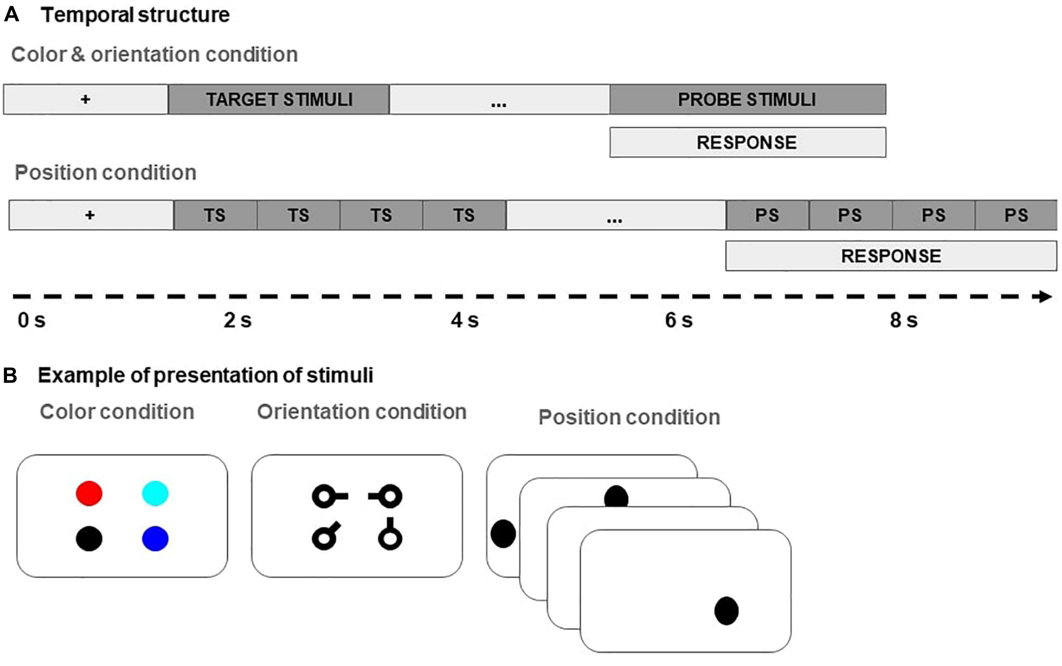


FIGURE 2 | (A) Timing structure; **(B)** stimuli of the experiment.

experience. The interviewer guided them through the empirical phenomenological interview in which participants' subjective reports were gathered.

The interview was designed to address the research question (i.e., how participants experience solving the visuo-spatial working memory task). It integrated interviewing approaches from qualitative research–specifically, constructivist grounded theory (Charmaz, 2004; Mills et al., 2006a,b; Charmaz and Belgrave, 2007)–and extant empirical phenomenological approaches, such as second-person in-depth phenomenological inquiry (Kordeš and Klauser, 2016), micro-phenomenological interview (Petitmengin, 2006), and descriptive experience sampling (Hurlburt, 2011).

Constructivist grounded theory and second-person in-depth phenomenological inquiry are both approaches that consider the constructive role of both the researcher and the participants in the process of knowledge-creation. The methods require participants to be highly engaged with the study. In turn, the interested position of the participants–unlike some contemporary approaches to first-person research (e.g., Petitmengin, 2006; Hurlburt, 2011)–allow the participants to go beyond an open-ended style of interviewing. Focused questions allowing for further theory-construction may be asked (cf. Charmaz, 2004).

The descriptive experience sampling technique is based on randomly sampling experiential episodes of individuals who are trained to observe and report their experience. We modified the typical descriptive experience sampling approach, which is based on sampling experience in an ecological setting, by inducing the target experience (i.e., solving a visuo-spatial working memory task) in a laboratory setting.

Initially, the interviewer asked about the overall temporal structure of the participants' experience: what were the salient events that took place in their awareness and what was their temporal succession relative to the behavioral and functional structure of the working memory task (i.e., presentation of the target stimuli/encoding, delay/maintenance, probe stimuli/recall and response). Then, the interviewer guided the participant back to the earliest moment they reported on and gathered its detailed description. The interviewer guided the participants away from general statements, descriptions of their beliefs about their experience, folk-psychological theories, and scientific (psychological and phenomenological) concepts (Petitmengin et al., 2019; Valenzuela-Moguillansky and Vásquez-Rosati, 2019). When descriptions of all salient experiences were grounded in bodily feelings, mental gestures, specific sensory modalities, and/or attitudes, the interviewer guided the participants to the next event. When the participant reported that there is nothing more to add regarding the trial, they were invited to return to the working memory task.

Audio recordings of the interviews were gathered using an Olympus WS-852 digital voice recorder.

Procedure

Participants were asked to take part in four study sessions, each lasting 60 min. The four sessions took place within 2 weeks[1]. Data were collected in an empty classroom with a table sectioned off for research purposes. During the task performance, the interviewer sat behind the participants.

During the first session, after signing the consent form, the participants were told how to perform the task. This included instructions for the visuo-spatial change detection task, as well as how to observe their experience. Prior to the beginning of each session, the interviewer provided a verbal description of both the working memory task and the interviewing protocol. It was emphasized that the participants should set aside their assumptions about the nature of their experience and cognition and focus only on the events and processes that occur in their awareness.

During each session, two task conditions were tested. One condition was concluded at the 30-min mark of the session. To avoid overwhelming participants with too many task parameters, the stimuli with simultaneous presentation (color and orientation) were presented at the first session. Within it, half of the participants started with the color condition and half with the orientation condition. Sequential presentation (position) was introduced in the first part of the second session. In the next session, conditions were randomized. When a new task condition was introduced, the participants were given a sample trial. The procedure was repeated until an hour elapsed.

A preliminary qualitative analysis of each session took place within 24 h after the interview. These analyses informed what follow-up questions were asked in subsequent interviews. In other words, data acquisition and initial data analysis were conducted in parallel. The process of parallel analysis allowed us to focus on the salient aspects of experience during the interview process itself and narrow our focus on the data directly addressing our research question (cf. Charmaz, 2004; Flick, 2009; Kordeš and Klauser, 2016).

Data Analysis

A total of 124 sessions were conducted across all the participants. All in all, we gathered 501 samples of experience (on average, 16.2 per participant). On average, participants received 4.0 prompts per session. A handful of participants expressed a higher degree of interest in exploring their experience. Sessions with them consisted of fewer samples that were explored in more detail.

Based on current standards in first-person research (Hurlburt, 2011; Valenzuela-Moguillansky and Vásquez-Rosati, 2019), two criteria for determining the validity of individual samples were followed: (a) whether the participants focused their description on a single trial; and (b) whether they described their lived experience rather than theories, assumptions, or beliefs) (e.g., statements such as "I simply saved it in my working memory." or "I stored it in my brain."). If a given sample did not reach these criteria, we eliminated the sampled experience from the analysis. Commonly, first sessions were eliminated entirely. 376

[1] First-person research is an evolving discipline in the context of cognitive science. Ongoing discussions regarding the methodological recommendations and guidelines are being conducted within the community. It is important to note that the present study was conducted between April and November 2018. At such time, a completely bottom–up approach was favored by the researchers in the first-person research community (for an epistemological discussion on the role of theory in first-person research, see Kordeš, 2016). Thus, the data acquired in this study were not constrained by questionnaires inquiring into, for instance, personality structure or the expression of mood disorders. Since examples of studies constraining these different types of instruments were published in the intervening years (Madeiros et al., 2021a) as well as discussions resolving the relevant epistemological issues (Madeiros et al., 2021b; Oblak, 2021b), it is now possible to put forward research design addressing these relationships.

samples (133 for the color condition, 134 for the orientation condition, and 109 for the position condition) of the experience of solving the change-detection task were considered valid and were analyzed further.

Analysis of the Interviews

Interview data were analyzed according to the principles of constructivist grounded theory. The main analytical instrument in constructivist grounded theory is coding; that is, the process of assigning more general descriptive tags to sections of raw data (see section "Coding"). Codes were grouped together based on their descriptive similarities, resulting in many experiential categories. The analysis yielded a large taxonomy of experiential categories (**Figure 3**).

Coding

During the analysis, the samples from all participants and across the three tasks were grouped together. Valid samples were coded according to the principles of constructivist grounded theory (Charmaz, 2004). Our goal was to determine a system of classification that would fully describe the key aspects of experience associated with the performance of a change-detection task. Initially, our focus was on the explicit ways of how individuals solve the change detection task – what we can broadly refer to as "strategies." However, as the data were coded inductively, other experiential categories that will be in detail presented below emerged through the process of analysis.

Following inductive coding, we employed relational coding, i.e., we constructed meaningful relationships between individual codes (Flick, 2009). Most importantly, we grouped them into higher-order categories based on descriptive similarities. Categories were constructed so that they remained stable across all participants (i.e., both the participants who received formal training in cognitive science and phenomenology, and the participants naive to those fields). Considering the large amount of phenomenal data acquired, we constructed a broad taxonomy of experiences associated with solving a change detection task (e.g., the most well-developed codes span five levels of abstraction). It is important to note here that the data we initially gathered were both broad and detailed (e.g., we gathered descriptions of how participants experienced their mental space taking shape in their consciousness). In relational coding, we narrowed our focus and constructed experiential categories that are explicitly related to engaging with a psychological task, and specifically, solving a visuo-spatial working memory task. We omitted those aspects of experience that are tangential to our research goal (e.g., experiencing the need to urinate). Finally, in relational coding, we mapped some of the categories induced from the data to extant concepts from psychological, phenomenological, and neuroscientific literature.

As mentioned in the section "Procedure," throughout the course of the study, data acquisition and analysis were performed in parallel (Flick, 2009). Based on insights gained during the analysis, we inquired about specific questions in more detail. Further, we were able to check whether certain experiential categories we had induced in earlier interviews were valid by asking the participants about them in subsequent interviews (Charmaz, 2004). This means that in parallel to the process of data acquisition, provisional categories were constructed. The validity of these categories was then checked against subjective reports in subsequent interviews.

We observed a rich continuum of experiences associated with solving the change-detection task. As such, the relational coding yielded a complex taxonomy (**Figure 3**) of experiential states, spanning five levels of coding (denoted with Roman numerals). On the lowest level of coding (Level I), the categories refer to the smallest degree of abstraction from the raw transcriptions of the interview. As such, they are – for the most part – theoretically unburdened. Moving upwards through the levels of coding, the experiential categories are grouped together based on both their structural (i.e., descriptive) similarities, as well as on working memory literature. For example, we introduced differences in coding that are based on working-memory processes (i.e., target stimuli/encoding, delay/maintenance, and the presentation of the probe stimuli/recall) that a particular aspect of experience is associated with.

Construction of the Codebook

The taxonomy of experiential categories was organized in an annotated codebook. Constructing a codebook (sometimes referred to as the coding manual; Kalinowski et al., 2010) is a standard procedure in qualitative research, and has been productively used in empirical phenomenological studies as well (e.g., Hurlburt and Heavey, 2006; Kordeš et al., 2019; Schwartzman et al., 2020). Codebook is a text in which each experiential category is described using (a) a name; (b) a definition; (c) logical relationships with other categories (i.e., which categories are superordinate or subordinate to each other; (d) representative examples; and (e) considerations (in which specific differences between similar categories are explicated and demonstrated with examples). The codebook serves three purposes. First, it provides a way of organizing the large amount of data acquired in the study. Second, it serves as a criterion of validity (i.e., a valid taxonomy of experiential categories yields a logically consistent codebook). Finally, the codebook provides a quick way for independent researchers to familiarize themselves with our coding taxonomy. The codebook is made available in its entirety in **Supplementary Material C**.

Determining the Validity of Coding

To ensure that we have reached conceptual depth (sometimes also referred to as saturation; Saunders et al., 2018), which is the point when we have gathered enough data for constructing a theory, we used the annotated codebook approach (Nelson, 2017). This approach supplements the codebook with a saturation grid (hence annotated codebook). A saturation grid is a tabulation of interviews listed along the horizontal (in our case, we listed individual participants to make the large amount of data tractable) and the codes listed along the vertical. The occurrence of each new code is marked in the appropriate cells. When in several subsequent interviews, no new codes can be induced from

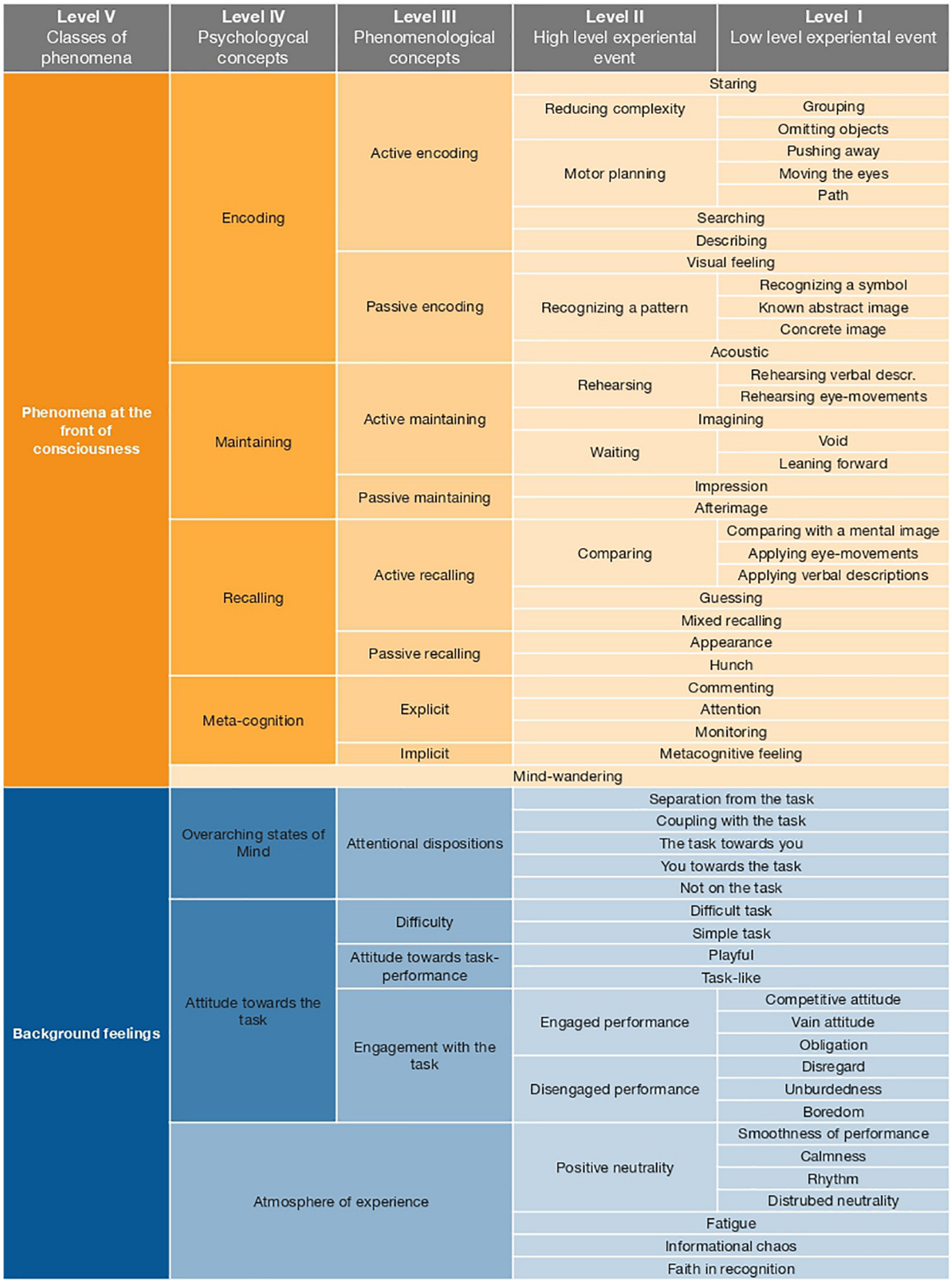


FIGURE 3 | Taxonomy of experiential categories.

TABLE 1 | Condensed saturation grid.

Code/Participant	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15 – 31
New codes discovered	22	4	0	8	5	1	2	0	0	2	0	0	0	1	0

the raw data, we can claim that we have reached conceptual depth. **Table 1** represents a condensed saturation grid for our study. We have reached saturation after completing the interviews with Participant 14.

To ensure the validity of our coding process we further took two measures: intercoder verification and consensual validation. The former refers to checking whether two independent coders reached the same types of codes on the same subset of data.

To this effect, the codebook was agreed on by the two principal coders, and the entirety of the gathered phenomenal data was subsequently subjected to the same codebook. Consensual validation refers to us checking with the participants themselves whether they agree that our coding process accurately reflects their experience. Consensual validation took the form of additional interviews in which the participants were given the experiential categories and asked to use them to report on their experience according to those categories. Participants were also given the opportunity to comment on whether the categories do not match their experience. No large changes were made by the participants except for remarking that the original name for the category impression should be changed from the Latin *imprimatura* to make it "less pretentious."[2] In total, five sessions were conducted with participants who were perceived by the researchers as most skilled in observing and reporting on their experience.

Epistemological Commitments

Since qualitative research, specifically phenomenology, may deviate from the positivist epistemology that is considered the standard fare in psychological research, this subsection explores the epistemological commitments that guided our study. Notably, many subjective phenomena (e.g., mind-wandering, cognitive strategies) that are commonly discussed in the literature, are phenomenologically opaque, since they are not described from the first-person perspective, but as third-person objects of inquiry (e.g., how they are operationalized in experimental research designs) (for a detailed discussion on this issue, see Varela et al., 1991; Varela, 1996; Thompson, 2007; Kordeš, 2016; Petitmengin, 2017).

Second, we adopt Hurlburt's (2011) notion of phenomenological data being radically non-subjective. By this, it is meant that phenomenal data that are presented in this study (a) refer to phenomena that are directly present to participants' consciousness (rather than general statements or recollections); (b) do not include participants' opinions and generalized statements on their experience. Nonetheless, the data are not objective because the only way to access them is through interviewing the individuals who lived through a certain experience.

Where our epistemology differs from Hurlburt's notion of radical non-subjectivity of phenomenal data is in the claim that there is an objectively correct way of describing experience. As noted above, we follow both the epistemology and methodology of constructivist grounded theory approach (Charmaz, 2004), which, broadly speaking, states that data are jointly constructed by the researchers and participants through the interviewing processes (Mills et al., 2006a,b). We attempted to increase the objectivity of the phenomenal data by: (a) conducting multiple interviewing sessions with the same participants; (b) gathering a larger amount of data than is typical for a qualitative study; and (c) engaging in the processes of consensual validation and intercoder verification.

See, however, the Section "Limitations and Future Directions" for how strategies and dispositions involved in working memory task performance might be constrained for a more objective understanding of the target phenomenology.

[2]It is to be noted that at the time of conducting this study, no procedures tailor-made for consensual validation of phenomenal data were available. The consensual validation presented in this study amounted to a first stage of developing a validation procedure for phenomenal data (for the development of the validation procedure, see Oblak, 2021a).

RESULTS

As mentioned in the Analysis section, we organized the qualitative data in the so-called annotated codebook. The experiential categories in the codebook span five levels of coding, with lower levels representing the smallest degree of abstraction from the raw data (i.e., amount to descriptions of experiential events as reported by the participants), whereas higher levels of coding refer to our attempts at organizing the data according to both descriptive similarities between categories, as well as insights from extant literature. The entire codebook is schematically represented in **Figure 3**.

At level V of coding – that is, the highest degree of abstraction from the raw data – we differentiate between phenomena at the front of consciousness and background feelings (cf. Colombetti, 2011). Phenomena at the front of consciousness refers to whatever is most present in the forefront of a participant's experience in each moment. These aspects of experience are readily available to conscious reflection even to participants who are not trained in observing and reporting on their experience. Conversely, background feelings describe overall, integrated, and more subtle aspects of experience that were not explicitly brought into the foreground of participants' awareness but nonetheless described how it was to be them when solving the visuo-spatial working memory task.

Phenomena at the Front of Consciousness

Phenomena at the front of consciousness consist of five broad categories: encoding, maintaining, recalling, meta-cognition, and mind-wandering. The categories encoding, maintaining, and recalling refer to the experience of the individual stage of the working memory task. Thus, these categories could broadly be referred to as mnemonic strategies, in that, they refer to the experience of attempting to solve the visuo-spatial working memory task. Each of these three categories is further subdivided into active (i.e., an aspect of experience perceived by participants as something they do) and passive (i.e., an aspect of experience perceived by participants as something that happens to them).

Active Encoding

Active encoding consists of (a) staring; (b) reducing complexity; (c) motor planning; (d) searching; and (e) describing. Staring refers to the experience of simply gazing at the target stimuli, hoping that you will somehow remember them. Often, the experience of staring is accompanied by the awareness that this strategy is in vain.

Reducing complexity is an experiential category where participants find the target stimuli to be too complex for them to be able to memorize them. For participants to be able to engage with the task, they first must simplify the stimuli for themselves. This is done in two ways, grouping, and omitting. Grouping refers to participants finding a commonality in some subset of stimuli (e.g., based on shape, color, common theme), whereby they can reduce the number of stimuli that have to be remembered. Conversely, omitting is an experiential category where participants – because of the complexity of the stimuli – consciously choose to remember only a subsection of stimuli, hoping that it will be enough for them to be able to recognize whether the probe stimuli are equal to or different from the target stimuli.

Motor planning refers to an experiential category that describes instances in which participants memorize the target stimuli by memorizing how they looked over them with their gaze. This category consists of three closely related but experientially different aspects of experience. Pushing away is the experience whereby individual stimuli feel as being weighed in space and moving across them with one's gaze is subjectively experienced as going against the apparent resistance of the stimuli. Moving the eyes is a strategy of memorizing the target stimuli whereby participants pay attention and remember the feeling in their eye muscles as they look over the stimuli. Finally, path refers to memorizing the trajectory of the center of the gaze (which may or may not be accompanied by the faint, bright line following it).

Searching is an experiential category that describes active encoding whereby participants attempt to discover a pattern in the stimuli, however, they need not actually find the pattern for this strategy to be successful. Finally, describing refers to participants tagging the target stimuli with some form of linguistic description, typically rendered in the form of inner speech.

Passive Encoding

Passive encoding refers to a set of experiences of the target stimuli whereby participants feel that memorization is an aspect of experience that happens to them. Passive encoding consists of three subcategories: (a) visual feeling; (b) recognizing a pattern; and (c) acoustics. Visual feeling is the experience of the overall visual atmosphere of the stimuli (e.g., blue, white, and black stimuli feeling cold).

Recognizing a pattern is an experiential category that describes the experience of some conceptual knowledge to describe the stimuli being immediately available upon perceiving them. Recognizing a pattern consists of three subcategories. First, recognizing a symbol refers to an awareness that there is some symbolic structure that can be readily related to the stimuli (e.g., white, green, and red being memorized as the colors of the Italian flag). Second, known abstract image refers to stimuli being describable through some geometrical shape (e.g., dots in the location condition outlining the shape of a star). Finally, concrete image describes instances where the stimuli elicit a mental image that incorporates some aspect of their appearance (e.g., in location condition, dots arranged in the shape of a rhombus might elicit a mental image of a kite under a blue sky).

The final category of passive encoding is acoustics. This aspect of experience refers to participants being aware of apparent sounds associated either with the appearance of an individual stage of the task or individual stimuli. Occasionally, this imaginary sound component may be sufficient for successful memorization.

Active Maintaining

Active maintaining describes participants willfully attempting to hold the target stimuli in their awareness during the delay period of the visuo-spatial working memory task. Active maintaining consists of three subcategories: (a) rehearsing; (b) imagining; and (c) waiting.

Rehearsing is an experiential category that refers to some aspect of experience being continuously repeated in a participant's awareness for them not to forget the target stimuli. Rehearsing consists of two subcategories: rehearsing a verbal description and rehearsing eye-movement (in which the method of active memorizing of motor planning is rehearsed).

Imagining describes the experience of the target stimuli being recreated in one's imagination using mental imagery (i.e., the target stimuli are maintained in the visual modality).

Finally, active maintaining consists of waiting. Interestingly, waiting was experienced as something that was actively performed by participants. It consists of two categories: void and leaning forward. The former describes an awareness of the absence of mental content. Participants are simply waiting for something to happen. Conversely, leaning forward describes the experience of performing a gesture toward the future moment (e.g., by attempting to anticipate or predict the identity of the probe stimuli).

Passive Maintaining

Passive maintaining is an experiential category that describes instances where the target stimuli remain in participants' awareness seemingly of their own accord. No explicit mental acts are required for these aspects of experience. Passive maintaining consists of two subcategories: (a) impression; and (b) afterimage.

Impression is an awareness of marked space left behind by the disappearing target stimuli during the delay period. It is a spatial feeling of something having been there. Conversely, afterimage is a visual experience in the form of a rapidly deteriorating echo of the target stimuli. Afterimage is typically the opposite (i.e., contrast) color to the stimuli.

Active Recalling

Active recalling is the experience of participants attempting to use an explicit strategy to determine whether the probe stimuli are equal to or different from the target stimuli. Three strategies of active recalling were identified: (a) comparing; (b) guessing; and (c) mixed recalling.

Comparing is an experience in which two explicit experiences at the foreground of one's consciousness are compared one against the other: first, there is some explicitly present memory of the target stimuli, and second, the visual presence of the probe stimuli. Three common experiences of comparing were observed: comparing with a mental image (i.e., a visually present memory of the target stimuli), applying eye-movements, and applying verbal descriptions.

Guessing refers to participants not having an explicitly present answer to the probe. Thus, they opt for picking an answer at random.

Finally, mixed recalling refers to a highly specific experiential dynamic during the probe stimuli. Namely, participants first passively become aware of whether the probe stimuli are equal to or different from the target stimuli (see below for passive recalling). In absence of an explicit strategy, participants feel unsure in their response. Thus, they use an additional explicit strategy to make sure that their initial feeling is correct.

Passive Recalling

Passive recalling is an experiential category that describes instances in which the knowledge about whether the probe stimuli are equal to or different from target stimuli occurs to the participants. Passive recalling contains two subcategories: (a) appearance; and (b) hunch. The two subcategories differ primarily in their location within one's experiential field, and the level of certainty. Appearance is an experiential category that describes the answer being present immediately in the perception of the probe stimuli. Appearance is accompanied by a sense of absolute certainty that the answer is correct. Second, hunch is a bodily felt likelihood regarding the answer to the probe stimuli (cf. gut feeling). Unlike appearance, hunch is subject to doubt.

Meta-Cognition

During the performance of the visuo-spatial working memory task, participants commonly reflect on their performance. We coded such experiences as meta-cognition. Subcategories of meta-cognition are divided along the lines of how saliently they are present in participants' awareness: (a) explicit; and (b) implicit. Three explicit aspects of meta-cognitive phenomenology were observed. First, commenting refers to an ongoing monolog or a dialogue, rendered in inner speech, providing feedback on what participants are doing and how well they are performing at the visuo-spatial working memory task. Second, attention refers to a sense of self-assuredness associated with the knowledge that they are paying close attention to the task. Since they are mindful of the task, the logic goes, it is unlikely that they missed something. Finally, monitoring refers to participants assuming a variety of observational perspectives of their experience of the task, both to make sure that there are no experiences that slip past them, and to be able to report on their experience.

The only implicit aspect of meta-cognitive phenomenology is meta-cognitive feeling. This experience is closely related to the experience of positive neutrality (described in detail below). Meta-cognitive feeling refers to whether participants suddenly become aware of some disturbance in their field of experience, signifying that they have made a mistake.

Mind-Wandering

The final phenomenon at the front of consciousness refers to mind-wandering. These are instances in which participants' attention moves away from the task toward other objects and/or mental activities, unrelated to the visuo-spatial working memory task, that they construct for themselves.

It is important to emphasize that many phenomena at the front of consciousness occur simultaneously or overlap within a single experiential episode. Phenomena at the front of consciousness are of particular interest to the theoretical construct of working memory, specifically, what kind of cognitive strategies, as accessible to conscious reflection, participants deploy to successfully solve the task. A detailed description of the categories and a quantitative analysis of experiential categories will be presented elsewhere. In the remainder of this paper, we pay particular attention to the second Level-V experiential category: background feelings.

Background Feelings

During data acquisition and analysis, we observed a huge variability in non-strategic experiences during the task-performance. While aspects of experience related to strategic performance of a visuo-spatial working memory task (outlined above) were - in a manner of speaking - artificial (being tied to specific stages of the working memory task), we observed an aspect of experience that provides a much more holistic description of how it feels to solve a visuo-spatial working memory task. These aspects of experience represent the second level-V category: Background feelings. Background feelings constitute aspects of experience that are not placed in the focus of participants' awareness, but nonetheless play a major role in the description of how it is to be someone in each moment. They are represented by several subtle attentional modulations and bodily feelings.

Notably, while phenomena at the front of consciousness are easily accessible even to participants who are not trained in observing one's experience, background feelings are not as readily apparent in conscious reflection. Background feelings include the following level IV subcategories: (a) attentional dispositions, (b) atmosphere of experience, and (c) attitude toward the task. These aspects of experience will be explored in detail in subsequent sections.

Attentional Dispositions

Attentional dispositions describe different attitudes we can take within our attention toward a specific object of our awareness. In turn, this attitude influences how we experience the object itself (cf. Petitmengin et al., 2009; Kordeš et al., 2019). In this study, we observed five different experiential categories that we consider to constitute attentional dispositions we may take toward a visual-spatial working memory task: separation from the task, coupling with the task, you toward the task, the task toward you, and not on the task. Separation from the task refers to an attitude of distance toward the change-detection task. Rather than being experienced as something that the participant has a causal effect over, the task is seen more as a video being passively observed. As Participant 27 reports:

> "I wasn't really with the task. I was solving it and it was important to me to solve it well, but I was also really aware of the room and you sitting behind me [...] What happened was that there was this sense of whole that was left over from the previous example, and

> then I didn't know whether I'm comparing it to this example or the one that came before." (VR.WM.1.27-04-O-01)

Conversely, coupling with the task is an experiential category that describes a spatially felt connection between the participant and the task at hand. This connection can be implicit and can appear simply as a loss of awareness of the research setting, the researcher, and the environment. Conversely, it can be very apparent, experienced as a sense of an enclosed, private space between the participant and the psychological task:

> "It is a kind of bubble. My attention is concentrated here at the front, at the task. This bubble is essentially a very pleasant kind of attention. I really like this focus. And my surroundings disappear. And I'm not aware of you either. There wasn't even an awareness of where I'm sitting. All that existed was this task. And I felt warm about that. This warmth was for me and for the way I do the task. It wasn't warm like the sun, but like a warmth in my thoughts." (VR.WM.1.24-01-P-02)

Alternatively, participants can approach the task by performing several mental gestures upon it (you toward the task). This experience may be so strong that it appears as a sense of direction moving from the participant toward the task. Consider the following example where the performance of mental gestures is explicitly reported on:

> "The task seemed more difficult than before. Before, it was enough for me to just have the impression of the image, whereas now I had to keep more of my focus on the task. I had to push away the things unrelated to the task. I performed movements with my gaze from object to object. I was jumping across them." (VR.WM.1.10-01-C-06)

Participants may also assume the attentional disposition of the task toward you. This experiential category refers to the participants appreciating the stimuli such as they are, without performing mental gestures upon it; that is, the category constitutes a receptive attitude toward the task. Consider the following report:

> "I just let the dots happen [...] I trusted myself that I will be able to know whether the next stimulus is correct or not. [...] It was as if I shut the task down. I was looking at the [fixation cross] and it was a little bit like meditating. I didn't have to think of the dots or anything." (VR.WM.1.15-02-P-04)

The final attentional disposition that we have observed within this study is not on the task, which refers to situations where participants attended to something other than the visual-spatial working memory task. This attentional disposition, however, is quite rare, as it was difficult to come across cases where the task did not in any way enter participants' awareness without the research context breaking down (i.e., without them stopping to solve the task).

Attitude Toward the Task

The second subcategory of background feelings are attitudes toward the task. These are background feelings within which the participants implicitly interpret the experimental paradigm in one way or another. In this context, we use the term "interpretation" to refer to what the participants make of the whole research setting. Attitudes toward the task contain three major ways of interpreting the change-detection task: (a) difficulty, (b) attitude toward task-performance, and (c) engagement with the task.

Difficulty is an attitude toward the task in which the participants interpret the task either as easy or difficult. Importantly, it is apparent that the difficulty of the task is not a property of an observer-independent, objective world, but it amounts to an attitude that participants take.

Attitude toward task-performance is an experiential category that refers to whether the visuo-spatial working memory task is appraised by participants as playful or task-like. Consider the following example, when the task appears playful:

> "The task seems more dynamic somehow. It was easier and more fun. I didn't know where the objects will appear and so there was always a little surprise when they showed up. I didn't have an anticipatory feeling about where they will appear. It was fun." (VR.WM.1.06-02-P-01)

Conversely, consider the following example when the experimental setup appears task-like:

> "I began thinking about how I don't like the exercise. The tails of the objects bothered me. They caused these bad feelings inside of me. These were more mental than bodily. It wasn't as if I was in pain or under stress. It was more of a preference. As if I didn't have these tails on the objects." (VR.WM.1.25-03-O-02)

Engagement with the task is an experiential category that describes the level of enthusiasm and zeal with which the participants approach it. In other words, it is a question of whether the participants approach the task as if it is important and something they must do well, or whether they approach it as something irrelevant, something that is simply there. Engagement with the task can be subdivided into engaged performance (where it is important to participants that they do well on the task), and disengaged performance (where the task is no longer at the center of their awareness). The former contains the following subcategories: (a) competitive attitude; (b) vain attitude; and (c) obligation. Disengaged performance contains the following subcategories: (a) disregard; (b) unburdedness; and (c) boredom.

The Atmosphere of Experience

The final subcategory of background feelings is the atmosphere of experience. It is a background feeling that describes the frame of the participant's experience. It refers to those aspects of experience that do not carry with them a specific content, but rather color whatever is at the center of the participant's experience. It contains four subcategories: (a) Positive neutrality, (b) fatigue, (c) informational chaos, and (d) faith in recognition. Fatigue is an experiential category that is to be understood in the trivial sense of the word – the experience of being tired.

Faith in recognition refers to an experience where despite not having an explicit mnemonic strategy, the participants feel that they will be able to successfully solve the change-detection task:

> "I know I can do it. It's just that I don't know how exactly I know. When I look at the memory stimulus, it just seems like I will be able to remember it." (VR.WM.1.07-01-C-02)

Informational chaos is a particular atmosphere of experience where the visual-spatial working memory task seems so complex as to be impossible to solve. It seems impossible to memorize the memory stimuli on account of how complex they seem. Participants commonly experience this complexity as the absence of any stable point at which to direct their attention to start memorizing the stimulus. It is an exceedingly unpleasant experience, and it is accompanied by feelings of anxiety and fear (cf. Kordeš, 2019). Over the course of the validation interview, Participant 22 reported that whenever she was solving the orientation condition, she was constantly anxious:

> "The memory stimulus appeared on the screen, and I immediately realized that I do not have the time to encompass the whole image with my eyes, let alone memorize it. I was moving my eyes around. I was panicking. I was trying to construct a memory, a pattern that would be sensible enough to be memorable. When the test stimulus actually happened, I had no explicit memory representation. There was no sense of how familiar the second stimulus was. So, I guessed the answer." (VR.WM.1.22-01-O-01)

Consider the further report from Participant 01 about the negative feelings incurred by the task:

> "It was terrible! I felt really unpleasant because I was fixated on having to get all of them right. I can feel the tension in my body, and I feel stressed. I have a feeling that this is how tasks are supposed to be performed. This is not research. It's an exam. It's something I have to get right. I know I have to get it right. This was constantly in the front of my awareness. And it took a lot of attention for me to be able to solve the task. I couldn't find any reference point that would help me remember it. I just waited for [the stimulus] to happen and try to remember its impression somehow. I was never sure about the answer. I always doubted the answer." (VR.WM.1.01-04-P-01)

The transcript of the interview does not do justice to the participant's distress. In research memos, we noted that her face became flushed, she was visibly uncomfortable, and bordering on tears. Participant 27 similarly reports:

> "It seemed to me like it should be fun, but it isn't. Now, it's stressful. I am using only about 20% to pay attention to it and I kind of don't care about it. I started to think about the task and now it's no longer as fun. And everything began going by so quickly. And so, I started thinking, why should I even bother if everything is going by so quickly. And then I ask myself, what's wrong with me? Am I dumb? And so, I just refuse to answer." (VR.WM.1.27-04-O-02)

One of the key aspects of informational chaos is the experience of the lack of stability, that is, the sense that there is no part of the visual experience that the participants could anchor their attention to in order to start remembering. Participant 25 reports:

> "I ran out of focus, but not because I was wrong with the previous one. It was simply no longer clear what I have to do. I don't understand what is happening on the screen. I don't know what to do. It was really, really uncomfortable. This panic was much worse than if I was just wrong. I almost started crying. I felt totally powerless. I felt this tension that was rising from my chest to my clavicles. I felt that I have an increased heart rate. I didn't know what to do. There was a confusion in my mind. I didn't know what to focus on. There was a lot of pondering, but actually, I didn't really have any thoughts. It was all mostly bodily experience. I couldn't stop it. I couldn't make a decision about what to do. This panic was the only thing I was really experiencing. I wasn't solving the task. I was just randomly clicking. Because I knew I had to. And there was this idea of a record of everything that I don't click on." (VR.WM.1.25-01-O-02)

She goes on:

> "I began thinking about how I don't like the exercise. The tails of the objects bothered me. They caused these bad feelings inside of me. These were more mental than bodily. It wasn't as if I was in pain or under stress. It was more of a preference. As if I didn't have these tails on the objects." (VR.WM.1.25-03-O-02)

Importantly, informational chaos is associated with a particular attitude toward the task, namely the sense of obligation. As Participant 22 reports:

> "I had a feeling of a total lack of control. I knew that I could not answer these questions. I randomly clicked. And when I finally gave into it, I had no thoughts about it, even though it encompassed the better part of my experience." (VR.WM.1.22-01-0-01)

The performative dimension of informational chaos is apparent: the participants are no longer solving the task, but are rather clicking at random, because of an experience of obligation. Rather than asking the researcher for the study to end – which they are told is a possibility at the beginning of the study – the participants continue to pretend to perform the working memory task.

Positive neutrality is an atmosphere of experience in which the participants are aware of the absence of any kind of disturbance. It is the sense of the smooth running of events. This smoothness; however, is not experienced positively in and of itself: it is a noticeably neutral aspect of experience. Because the participants are aware of nothing being wrong, this neutrality is experienced as something good, as something mildly positive. Positive neutrality may be brought into one's awareness four interrelated and connected ways: (a) Smoothness of performance; (b) calmness; and (c) rhythm.

Smoothness of performance is an atmosphere of experience in which the participants are aware of positive neutrality at their own task performance. The task – and its performance – run smoothly, without any interruptions. This smoothness in and of itself is not pleasant. The absence of disturbances is the aspect of experience that is perceived as positive. Consider the following:

> "Everything was related to the rhythm. It created this flow. This rhythm was silent, discrete, a staccato. I moved my eyes according to where I was expecting the dot to appear. When it didn't appear, I got a sense of dissonance. It went outside of the rhythm. It's a very mental experience, yes, but it is not reasoning. It wasn't as if I said 'Aha, it went out of rhythm, therefore, it does not fit.' It was a more subtle feeling. Like an aha moment. The dot was not here, so it must be there. When the dot was in an unexpected place,

> the rhythm broke down. And so did any image that I had in my mind." (VR.WM.1.19-02-P-02)

The experience of rhythm is, by and large, tied to the appearance of stimuli themselves. This experiential category refers to an embodied (and sometimes even auditory, when accompanied by the experience of acoustics) element of experience. Participants report that rhythm structures their time and that task-performance in consequence becomes more relaxed and predictable. Rhythm can be embodied and is commonly tied to the feeling of heart rate. In a similar manner to the smoothness of performance, rhythm can break down in the case of a mistake or a disturbance in an experiential field. This leads to the tempo stopping. Insofar as the disturbance does not occur, the rhythm encompasses the final key press as well:

> "A rhythm of performing the task established itself. At the end, I answer completely automatically. If there's a mismatch, this way of solving the task stops." (VR.WM.1.17-04-O-02)

Smoothness of performance is an experience that is structurally like rhythm from the point of view that once it is present, it appears as the default state of the experiential landscape. Both aspects of experience may be cut short by a disturbance. The difference is that smoothness of performance is, by and large, unconscious. Rhythm, on the other hand, clearly constitutes a presence. Consider the following example:

> "I said to myself, 'one, two, three, four.' This happened in the same rhythm as the dots appeared in. these verbalizations were sounds that happened in my own voice. The rhythm helped me concentrate. It gave structure to my sense of time. It wasn't intended to be a dancing rhythm but it reminded me a lot about dance where you have to count [...] I'm the author of these words. I willfully produce them. When they first appeared, I didn't willfully create them. It wasn't my strategy at first. But then I noticed that I'm doing it and that it feels good. What felt good was the structure. I latched on to that and started actively doing it." (VR.WM.1.04-02-P-01)

Finally, calmness is an atmosphere of experience whereby the participants feel pleasant in the absence of disturbances. Calmness itself has no object. It describes the general way of existing in a particular moment.

DISCUSSION

In the present study, we investigated the experience during solving a visuo-spatial working memory task. We used the change-detection task, a standard instrument for studying working memory, and gathered first-person data with a combination of experience sampling and an in-depth phenomenological interview. This resulted in a large array of data depicting the experience of participants solving the visuo-spatial working memory task. During the study, we observed that the experiential landscapes associated with visuo-spatial working memory task-performance are extremely rich, both in terms of mental acts that participants use to attempt to "solve" the task, as well as in terms of depth of experience; that is, overall experiential states accompanying (and often probably determining) the task-performance.

Our qualitative analysis separated the gathered phenomenal data into two major classes of experience. First, there is an experiential dimension, coded as phenomena at the front of consciousness. This category describes aspects of experience that occupy the center of participants' awareness and are readily accessible to consciousness reflection even in absence of training in how to observe and report on experience. Phenomena at the front of consciousness consist of various strategies of solving the visuo-spatial working memory task, together with meta-cognitive experiences, and mind-wandering.

Second, during the analysis, an experiential dimension that is harder to access in conscious reflection as it comprises the overall sense of experience (e.g., bodily feelings, emotional atmosphere, mood) was consistently detected. We coded this experiential dimension as background feelings. Most of the present article, including the discussion, is dedicated to background feelings. The study of background feelings poses significant methodological challenges, but – we suspect – could play an important role in understanding and interpreting the results of visuo-spatial working memory task-performance, as well as working memory more broadly.

Comparison to Existing Ideas

In the present subsection, we tie the experiential categories observed in this study with extant ideas from mind sciences. The identified categories that belong to phenomena at the front of consciousness map well onto extant constructs from the domains of cognitive strategies, metacognition, and perception. The broad distinction between active and passive experiential categories fits the discussion on agency in philosophy of mind (Gallagher, 2012). A number of categories were observed that might be summarized as representational modalities (imagining, motor planning, describing) that correspond with the basic types of experiences identified via descriptive experience sampling (inner seeing, inner speech, feeling) (Heavey and Hurlburt, 2006). Reducing complexity can be linked to chunking (Miller, 1956). The complex phenomenological dynamic seen during the delay period (i.e., the present moment can be experienced as absence of experience or an anticipation of the future) can be linked to experimental results that suggest that encoding and recall are more akin to different styles of solving the task than separate memory mechanisms (Speer et al., 2003). Finally, the embodied dimension of hunch suggests that it may be integrated into the somatic marker hypothesis, a theory that suggests intuition is the statistical awareness of the probability of some event, reflected in the personal level as a specific array of bodily feelings (Damasio, 1996). A more detailed examination of phenomena at the front of consciousness will be presented in a follow-up publication.

While phenomena at the front of consciousness amount to well-known constructs within the sciences of the mind, until recently, background feelings received little attention within psychological literature (although see Tsuchiya and Koch, 2016; de Haan, 2020). Yet, we believe that understanding this aspect of experience may be as important as understanding the strategies participants use to solve a visuo-spatial working memory task for disentangling working memory task performance.

By way of example, let us examine one of the categories that we identified in this study: positive neutrality, a subcategory of the background feeling atmosphere of experience. Positive neutrality refers to a subtle awareness of an absence of change. Participants report solving the visuo-spatial working memory task guided by the awareness of an absence of disturbances (e.g., wrong answers). This lack of disturbances itself is felt as neutral. However, since - in the context of visuo-spatial working memory task performance – it signifies that the participants have not detected any blunders, it is appraised as slightly positive. In such cases, positive neutrality also reflects the repeating, rhythmic pattern of trials occurring at a predictable rate. Only when something goes wrong (e.g., the pace at which they are scanning the stimuli, wrong answer, end of task, etc.) is this positive neutrality disturbed.

These experiential dynamics mirrors the notion of surprise in the so-called free energy principle theories of cognition (Friston, 2009, 2010). According to this theory, living organisms work to minimize their internal disorder (entropy). They create a barrier between themselves and the outside world, wherein inside of this boundary, they maintain their own homeostasis (Kirchhoff et al., 2018; Demekas et al., 2020). A surprising event in this framework, is an event that was not predicted by the organism (Schwartenbeck et al., 2013). When applied to cognition, free energy principle is related to predictive processing, the notion that to minimize the computational complexity of incoming sensory information, organisms predict the likeliest situation and then only process surprising stimuli (i.e., stimuli that they were unable to predict) (Seth et al., 2012; Seth, 2014; Clark, 2016).

The properties of some of the background feelings (e.g., the feeling of positive neutrality) are remarkably like phenomenological properties associated with or allowing for the experience of surprise (Bitbol, 2019). Integrating these ideas with our own, we posit that positive neutrality is the stable state of successful prediction. When positive neutrality is disturbed, the experiencing individual is prompted to acknowledge change.

Second, the phenomenon of attentional dispositions has been commonly reported in empirical phenomenological literature. Originally, it was reported in the study on how we experience attending to sound (Petitmengin et al., 2009). In that study, attentional dispositions refer to the changes in our experience of the source of sound when we attend to it with different attitudes (e.g., as a physical sensation in our ears, as a location in space, as an acoustic experience with volume, pitch, and timbre). The findings were corroborated in a study on the experience of meditation (Kordeš et al., 2019). Specifically, participants reported experiencing pain (e.g., in their legs or back when meditating). Pain could then be attended to with different attitudes, which, in turn, altered the intensity of noxious sensations.

Some attentional dispositions bear resemblance to what Oblak et al. (2021) refer to as affective resonance, a property of objects present in one's environment that they can jointly form higher-order systems (see also Thompson, 2007; Kyselo and Tschacher, 2014). This description corresponds to the disposition of coupling with the task, where participants report feeling that the task-performance is not something that they do themselves, but they do it in concert with the computer. Conversely, the disposition separation from the task is descriptively similar to the phenomenological notion of detunement. Detunement refers to an inability of accessing socially shared space in depression (Fuchs, 2005, 2007), schizophrenia (Krueger and Aiken, 2016; Silverstein et al., 2017), and anxiety (Trigg, 2017). It may be that part of why individuals with specific psychopathologies underperform on working memory tasks (Christopher and MacDonald, 2005; Rose and Ebmeier, 2006; Forbes et al., 2009; Moran, 2016) is because phenomenologically, the task are not salient enough in their awareness for them to be able to attentionally engage with them.

In line with Hurlburt et al.'s (2016) findings, some insights, associated with background feelings, indicate that the experience of visuo-spatial working memory task-performance are quite removed from day-to-day experiences. Participant 22, for example, reports feelings (coded as informational chaos) that are familiar to her only from other times when she was solving similar psychological tasks:

> "I remember solving a similar task for [one of the professors]. I was stuck in some lab, I was a little bit afraid, and on top of that, there was this bombardment of stimuli that was totally overwhelming. I really don't recognize this feeling from situations other than stuff like this." (VR.WM.1.22-01-O-01)

Gathered phenomenal data raise questions about the validity of studies operating under the assumption that psychological tasks elicit only the target phenomenon without checking participant's experience of said task (cf. Hurlburt, 2011, Chapter 21). Our study has not gathered sufficient data to conclusively reach this conclusion. Additional investigations of experience underlying performance of visuo-spatial working memory tasks, as well as more ecological (e.g., ethnographic) research of day-to-day experience associated with visuo-spatial working memory are needed. For example, recent qualitative studies into the experience of working memory tasks have demonstrated the presence of various emotional states (bearing similarities to our experiential category informational chaos), primarily dealing with the social expectations and that are associated with performance accuracy (Laybourn et al., 2022). Thus, this study points to some as of yet unanswered questions in the broader field of working memory research.

Toward a Neurophenomenology of Working Memory

Positive neutrality and attentional dispositions are not the only example of an experiential category where we observed background feelings playing an important (and perhaps essential) role in task-performance. The entire group of experiential categories we coded as the atmosphere of experience was – according to subjective estimates of participants – important when engaging with the visuo-spatial working memory task. Our results seem to coincide with some recent ideas in cognitive science indicating the importance of understanding the general, background atmospheres. One of such ideas is the theory of overarching states of mind (Herz et al., 2020). It proposes that, on the level of neural dynamics, we can observe integrated states that coordinate all other aspects of cognition, such as perception, emotions, thinking, as well as embodied behavior. It is not unreasonable to hypothesize that such overarching states of mind can be detected on the personal level of description as a specific class of background feelings.

We hope to have shown that aspects of experience that are not so easily identifiable form a significant part of an individual's experiential landscape. Detection of such aspects of experience requires participants who, at the very least, are interested in exploring their own experience, but preferably are also trained in observing and reporting their experience.

Our results correspond to various observations by many within the field of first-person research (Petitmengin, 2006; Hurlburt, 2011; Kordeš, 2016), thereby indicating the necessity of re-examining the traditional roles of researcher and participant (cf. Orne, 1962) in the domain of researching experience. Since the participant is the only person with the access to experience under investigation, then, her process or reflection constitutes the principal instrument of inquiry. If this instrument is inadequate, then, no subsequent step (e.g., interview, analysis, etc.), no matter how refined, can improve the validity of a study. Thus, it is necessary to consider the people whose experience we are investigating not as passive subjects from whom data can be objectively extracted, but rather as equal partners in the research process – as co-researchers (cf. Kordeš and Klauser, 2016).

The acquired first-person data allow us to put forward a conjecture. It may be that the aspects of experience detected in this study (and coded as background feelings) determine our behavior to a greater extent than the readily accessible phenomena at the forefront of consciousness. It may be that a larger part of visuo-spatial working memory task-performance occurs below the threshold of conscious awareness (cf. Soto et al., 2011; Dutta et al., 2014; Soto and Silvanto, 2014). Explicit, conscious attempts at taking control over the task-performance (i.e., assuming a strategic attitude) may therefore only serve to interfere with task-performance. In other words, attempts at explicit performance stop subconscious task-performance (cf. Lyubomirsky et al., 1999).

However, when talking about a large part of task-performance taking place below the threshold of conscious awareness, we are referring to an everyday, and, for the most part, inaccurate awareness. Is it possible that a more precise insight into the content of consciousness allows this threshold to be raised? What if a more detailed reflection can be used to detect traces of cognitive processes taking part in visuo-spatial working memory task-performance? We believe that this is the case and that these traces can be observed, but only by participants, skilled enough to be able to access background feelings (cf. Petitmengin et al., 2007). Thus, it is essential for the validity of studies attempting to provide an in-depth account of first-person data to use participants who are skilled and trained in observing and reporting on their experience (cf. Miyahara et al., 2020).

In future research, we aim at replicating the study presented in this paper, however, recruiting only participants trained in observing and reporting on their experience. Additionally, we hold that using trained participants may be beneficial for future fMRI and neurophenomenological studies (cf. Lutz et al., 2002; Fernyhough et al., 2018). As demonstrated by Hurlburt et al. (2016), validity of studies assuming that the experimental paradigm constructs the phenomenon of inquiry, without verifying it on the level of experience, is questionable (Hurlburt, 2011).

LIMITATIONS AND FUTURE DIRECTIONS

Although in addition to first-person reports we also collected behavioral responses (i.e., performance accuracy and reaction times), this paper omitted quantitative analyses of the data. There are two reasons for this. First, as the aim of this study – i.e., exploring the breadth of experiences evoked by and accompanying performance of a visuo-spatial working memory task – was complex as it is, we wanted to avoid addressing additional research questions related to behavioral performance. For example, it would be interesting to explore the relationship between the identified experiential categories, behavioral measures (reaction times and performance accuracy) and task modality, however, this aspect of the results will be presented in the follow-up paper.

Second, the study followed the golden standard for estimating when enough qualitative data were gathered (see section "Determining the Validity of Coding"). However, despite gathering a large amount of data for a qualitative study, there are too few acquired samples to be able to draw any statistical conclusions.

Future studies should be designed specifically to constrain first-person data (i.e., reports on strategies and dispositions involved in solving a visuo-spatial working memory task), behavioral (performance accuracy and reaction times), and psychometric measures (e.g., personality questionnaires, clinical scales). Such a research design would necessitate a development of a custom-made framework for reporting on the phenomenology of solving a visuo-spatial working memory task in real time. A precedence for such an approach can be seen in the studies on the experience of inner speech. These began with sampling of overall experience (Heavey and Hurlburt, 2006), moving on to detailed account of inner speech (Hurlburt et al., 2013), before developing phenomenology-inspired questionnaires (Alderson-Day et al., 2018) and neuroimaging studies (Hurlburt et al., 2016).

CONCLUSION

In this paper, we presented an empirical phenomenological study of what individuals experience during solving a visuo-spatial working memory task (specifically, the change-detection task). Using a combination of experience sampling and an in-depth interview we gathered a large amount of experiential data. The data were analyzed according to the principles of constructivist grounded theory. The resulting experiential data demonstrate the wealth of different experiences associated with solving a working memory task. Some of them are easy-to-detect and could be productively used to further approaches within experimental cognitive psychology. Others amount to overarching, integrated states of consciousness within a given moment. These require further neurophenomenological investigation, as well as reconsidering the social dynamics between researchers and participants. Gathered phenomenal data raise questions about the validity of studies operating under the assumption that psychological tasks elicit only the target phenomenon without checking participant's experience of said task.

DATA AVAILABILITY STATEMENT

The original contributions presented in the study are publicly available. This data can be found here: https://osf.io/mkp8b/.

ETHICS STATEMENT

The studies involving human participants were reviewed and approved by the Internal Review Board at the University Psychiatric Clinic Ljubljana, University of Ljubljana. The patients/participants provided their written informed consent to participate in this study.

AUTHOR CONTRIBUTIONS

GR, ASO, and UK: conceptualization. AO, GR, ASO, and UK: methodology, and writing – review and editing. GR and ASO: software. AO and UK: formal analysis. AO: investigation and writing – original draft and project administration. AO and ASO: visualization. UK and GR: supervision and funding acquisition. All authors: contributed to the article and approved the submitted version.

FUNDING

This work was supported by Slovenian Research Agency (ARRS) research grants P5-0126 (UK), and J3-9264 and P3-0338 (ASO and GR).

ACKNOWLEDGMENTS

The authors wish to thank Ema Demšar for her help with conducting the study, and Jaya Caporusso, Asena Boyadzhieva, and Maša Rebernik for their help with preparing the manuscript.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fpsyg.2022.811712/full#supplementary-material

REFERENCES

Alderson-Day, B., Mitrenga, K., Wilkinson, S., McCarthy-Jones, S., and Fernyhough, C. (2018). The varieties of inner speech questionnaire – Revised (VISQ-R): replicating and refining links between inner speech and psychopathology. *Conscious. Cogn.* 65, 48–58. doi: 10.1016/j.concog.2018.07.001

Baddeley, A. D. (2010). Working memory. *Curr. Biol.* 20, R136–R140. doi: 10.1016/j.cub.2009.12.014

Baddeley, A. D., and Hitch, G. (1974). "Working memory," in *The Psychology of Learning and Motivation: Advances in Research and Theory*, ed. G. H. Bower (New York, NY: Academic Press), 47–89.

Barrouillet, P., Bernardin, S., Portrat, S., Vergauwe, E., and Camos, V. (2007). Time and cognitive load in working memory. *J. Exp. Psychol. Learn. Mem. Cogn.* 33, 570–585. doi: 10.1037/0278-7393.33.3.570

Bitbol, M. (2019). "Neurophenomenology of surprise: sparks of awakening," in *Surprise at the Intersection of Phenomenology and Linguistics*, eds N. Depraz and A. Celle (Amsterdam: John Benjamins), 10–21. doi: 10.1075/ceb.11.01bit

Bo, J., and Seidler, R. D. (2009). Visuospatial working memory capacity predicts the organization of acquired explicit motor sequences. *J. Neurophysiol.* 101, 3116–3125. doi: 10.1152/jn.00006.2009

Braga, L. W., Amemiya, E., Tauil, A., Suguieda, D., Lacerda, C., Klein, E., et al. (2017). Tracking adult literacy acquisition with functional MRI: a single-case study. *Brain Mind Educ.* 11, 1121–1132. doi: 10.1111/mbe.12143

Buchsbaum, B. R. (2013). The role of consciousness in the phonological loop: hidden in plain sight. *Front. Psychol.* 4:496. doi: 10.3389/fpsyg.2013.00496

Caramazza, A. (1986). On drawing inferences about the structure of normal cognitive systems from the analysis of patterns of impaired performance: the case for single-patient studies. *Brain Cogn.* 5, 41–66. doi: 10.1016/0278-2626(86)90061-8

Charmaz, K. (2004). *Constructing Grounded Theory: A Practical Guide Through Qualitative Analysis*. London: Sage Publications.

Charmaz, K., and Belgrave, L. L. (2007). *Grounded Theory. The Blackwell Encyclopedia of Sociology*. Malden, MA: Blackwell Pub.

Christopher, G., and MacDonald, J. (2005). The impact of clinical depression on working memory. *Cogn. Neuropsychiatry* 10, 379–399. doi: 10.1080/13546800444000128

Clark, A. (2016). *Surfing Uncertainty: Prediction, Action, and the Embodied Mind*. Oxford: Oxford University Press.

Colombetti, G. (2011). Varieties of pre-reflective self-awareness: foreground and background bodily feelings in emotion experience. *Inquiry Interdisciplinary J. Philos.* 54, 293–313. doi: 10.1080/0020174x.2011.575003

Coltheart, M. (2017). The assumptions of cognitive neuropsychology: reflections on Caramazza (1984)Caramazza (1986). *Cogn. Neuropsychol.* 34, 397–402. doi: 10.1080/02643294.2017.1324950

Coltheart, M., Curtis, B., Atkins, P., and Haller, M. (1993). Models of reading aloud: dual-route and parallel-distributed-processing approaches. *Psychol. Rev.* 100, 589–608. doi: 10.1037/0033-295X.100.4.589

Cummine, J., Gould, L., Zhou, C., Hrybouski, S., Siddiqi, Z., Chouinard, B., et al. (2013). Manipulating instructions strategically affects reliance on the ventral-lexical reading stream: converging evidence from neuroimaging and reaction time. *Brain Lang.* 125, 203–214. doi: 10.1016/j.bandl.2012.04.009

Cuthbert, B., and Insel, T. (2010). The data of diagnosis: new approaches to psychiatric classification. *Psychiatry* 73, 311–314. doi: 10.1521/psyc.2010.73.4.311

Damasio, A. R. (1996). The somatic marker hypothesis and the possible functions of the prefrontal cortex. *Philos. Trans. R. Soc. B.* 351:1346. doi: 10.1098/rstb.1996.0125

D'Angiulli, A., and Smith LeBeau, L. (2002). On boredom and experimentation in humans. *Ethics Behav.* 12, 167–176. doi: 10.1207/S15327019EB1202_4

de Haan, S. (2020). *Enactive Psychiatry*. Cambridge: Cambridge University Press.

De Hollander, G., Forstmann, B. U., and Brown, S. D. (2016). Different ways of linking behavioral and neural data via computational cognitive models. *Biol. Psychiatry Cogn. Neurosci. Neuroimag.* 1, 101–109. doi: 10.1016/j.bpsc.2015.11.004

Demekas, D., Parr, T., and Friston, K. J. (2020). An investigation of the free energy principle for emotion recognition. *Front. Computat. Neurosci.* 14:30. doi: 10.3389/fncom.2020.00030

Dutta, A., Shah, K., Silvanto, J., and Soto, D. (2014). Neural basis of non-conscious visual working memory. *NeuroImage* 91, 336–343. doi: 10.1016/j.neuroimage.2014.01.016

Fernyhough, C., Alderson-Day, B., Hurlburt, R. T., and Kühn, S. (2018). Investigating multiple streams of consciousness: using descriptive experience sampling to explore internally and externally directed streams of thought. *Front. Hum. Neurosci.* 12:494. doi: 10.3389/fnhum.2018.00494

Flannagan, O. (1992). *Consciousness Reconsidered*. Cambridge: The MIT Press.

Flick, U. (2009). *An Introduction to Qualitative Analysis*. London: Sage Publications.

Forbes, N. F., Carrick, L. A., McIntosh, A. M., and Lawrie, S. M. (2009). Working memory in schizophrenia: a meta-analysis. *Psychol. Med.* 39, 889–905. doi: 10.1017/s0033291708004558

Friston, K. (2009). The free-energy principle: a rough guide to the brain? *Trends Cogn. Sci.* 13, 293–301. doi: 10.1016/j.tics.2009.04.005

Friston, K. (2010). The free-energy principle: a unified brain theory? *Nat. Rev. Neurosci.* 11, 127–138. doi: 10.1038/nrn2787

Froese, T., Gould, C., and Barrett, A. (2011a). Re-viewing from within: a commentary on the first- and second-person methods in the science of consciousness. *Constructivist Foundations* 6, 254–269.

Froese, T., Gould, C., and Seth, A. K. (2011b). Validating and calibrating first- and second-person methods in the science of consciousness. *J. Conscious. Stud.* 18, 38–64.

Froese, T., Suzuki, K., Ogai, Y., and Ikegami, T. (2012). Using human–computer interfaces to investigate 'Mind-As-It-Could-Be' from the first-person perspective. *Cogn. Comput.* 4, 365–382. doi: 10.1007/s12559-012-9153-4

Fuchs, T. (2005). The phenomenology of body space and time in depression. *Comprendre* 15, 108–121.

Fuchs, T. (2007). Psychotherapy of the lived space: a phenomenological and ecological concept. *Am. J. Psychother.* 61, 423–439.

Gallagher, S. (2012). Multiple aspects of the sense of agency. *New Ideas Psychol.* 30, 15–31. doi: 10.1016/j.newideapsych.2010.03.003

Gallagher, S., Reinerman, L., Janz, B., Backelman, P., and Trempler, J. (2015). *A Neurophenomenology of Awe and Wonder: Towards a Non-reductionist Cognitive Science*. New York, NY: Palgrave Macmillan.

Heavey, C. L., and Hurlburt, R. T. (2006). The phenomena of inner experience. *Conscious. Cogn.* 17, 798–810. doi: 10.1016/j.concog.2007.12.006

Herz, N., Baror, S., and Bar, M. (2020). Overarching states of mind. *Trends Cogn. Sci.* 24, 184–199. doi: 10.1016/j.tics.2019.12.015

Hurlburt, R. (1990). *Sampling Normal and Schizophrenic Inner Experience*. New York, NY: Plenum Press.

Hurlburt, R. T. (2011). *Investigating Pristine Inner Experience: Moments of Truth*. Cambridge: The MIT Press.

Hurlburt, R. T., Alderson-Day, B., Kühn, S., and Fernyhough, C. (2016). Exploring the ecological validity of thinking on demand: neural correlates of elicited vs. spontaneously occurring inner speech. *PLoS One* 11:e0147932. doi: 10.1371/journal.pone.0147932

Hurlburt, R. T., and Heavey, C. L. (2006). *Exploring Inner Experience: The Descriptive Experience Sampling Method*. Amsterdam: John Benjamins.

Hurlburt, R. T., Heavey, C. L., and Kelsey, J. M. (2013). Toward a phenomenology of inner speaking. *Conscious Cogn.* 22, 1477–1494. doi: 10.1016/j.concog.2013.10.003

Ikeda, M., Iwanaga, M., and Seiwa, H. (1996). Test anxiety and working memory system. *Perceptual Motor Skills* 82, 1223–1231. doi: 10.2466/pms.1996.82.3c.1223

Ioannidis, J. P. (2005). Why most published research findings are false. *PLoS Med.* 2:e124. doi: 10.1371/journal.pmed.0020124

James, W. (1890). *The Principles of Psychology*. New York, NY: Holt.

Jaspers, K. (1997). *General Psychopathology*, Vol. 1. Baltimore: Johns Hopkins University Press.

Kalinowski, P., Lai, J., Fidler, F., and Cumming, G. (2010). Qualitative research: an essential part of statistical cognition research. *Stat. Educ. Res. J.* 9, 22–34. doi: 10.52041/serj.v9i2.373

Kirchhoff, M., Parr, T., Palacios, E., Friston, K., and Kiverstein, J. (2018). The Markov blankets of life: autonomy, active inference and the free energy principle. *J. R. Soc. Interface* 15:20170792. doi: 10.1098/rsif.2017.0792

König, S. U., Goeke, C., Meilinger, T., and König, P. (2017). Are allocentric spatial reference frames compatible with theories of enactivism? *Psychol. Res.* 83, 498–513. doi: 10.1007/s00426-017-0899-x

Kordeš, U. (2016). Going beyond theory: constructivism and empirical phenomenology. *Constructivist Foundations* 14:149152.

Kordeš, U. (2019). Horizons of analysis. *Constructivist Foundations* 14, 149–152.

Kordeš, U., and Klauser, F. (2016). Second-person in-depth phenomenological inquiry as an approach for studying enaction of beliefs. *Interdisciplinary Description Complex Syst.* 14, 369–377. doi: 10.7906/indecs.14.4.5

Kordeš, U., Oblak, A., Smrdu, M., and Demšar, E. (2019). Ethnography of meditation: an account of pursuing meditative practice as a tool for researching consciousness. *J. Conscious. Stud.* 26, 184–237.

Krueger, J., and Aiken, A. T. (2016). "Losing social space: phenomenological disruptions of spatiality and embodiment in Moebius Syndrome and Schizophrenia," in *Phenomenology and Science*, eds J. Reynolds and R. Sebold (London: Palgrave Macmillan), 121–139. doi: 10.1057/978-1-137-51605-3_7

Kyselo, M., and Tschacher, W. (2014). An enactive and dynamical systems theory account of dyadic relationships. *Front. Psychol.* 5:452. doi: 10.3389/fpsyg.2014.00452

Laybourn, S., Frenzel, A. C., Constant, M., and Liesefeld, H. R. (2022). Unintended emotions in the laboratory: emotions incidentally induced by a standard visual working memory task relate to task performance. *J. Exp. Psychol. Gen.* [Online ahead of print] doi: 10.1037/xge0001147

Lutz, A., Lachaux, J., Martinerie, J., and Varela, F. J. (2002). Guiding the study of brain dynamics by using first-person data: synchrony patterns correlate with ongoing conscious states during a simple visual task. *Proc. Natl. Acad. Sci. U.S.A.* 99, 1586–1591. doi: 10.1073/pnas.032658199

Lyubomirsky, S., Tucker, K. L., Caldwell, N. D., and Berg, K. (1999). Why ruminators are poor problem solvers: clues from the phenomenology of dysphoric rumination. *J. Pers. Soc. Psychol.* 77, 1041–1060. doi: 10.1037//0022-3514.77.5.1041

Madeiros, S., Crempien, C., Vásquez-Rosati, A., Duarte, J., Andreu, C., Langer, A. I., et al. (2021a). Assessing subjective processes and vulnerability in mindfulness-based interventions: a mixed methods exploratory study. *Constructivist Foundations* 16, 203–220.

Madeiros, S., Crempien, C., Vásquez-Rosati, A., Duarte, J., and Langer, A. I. (2021b). Authors' response: the art and science of befriending inner experience. *Constructivist Foundations* 16, 238–243.

Mcintyre, R. (1999). "Naturalizing phenomenology: dretske on qualia," in *Naturalizingphenomenology: Issues in Contemporaryphenomenology and Cognitive Science*, eds J. Petitot, F. J. Varela, B. Pachoud, and J. Roy (Stanford: Stanford University Press), 429–439. doi: 10.1515/9781503617421-018

Miller, G. A. (1956). The magical number seven, plus or minus two: some limits on our capacity for processing information. *Psychol. Rev.* 63, 81–97. doi: 10.1037/h0043158

Miller, M. B., Donovan, C.-L., Bennett, C. M., Aminoff, E. M., and Mayer, R. E. (2012). Individual differences in cognitive style and strategy predict similarities in the patterns of brain activity between individuals. *NeuroImage* 59, 83–93. doi: 10.1016/j.neuroimage.2011.05.060

Mills, J., Bonner, A., and Francis, K. (2006a). Adopting a constructivist approach to grounded theory: implications for design. *Int. J. Nurs. Pract.* 12, 8–13. doi: 10.1111/j.1440-172X.2006.00543.x

Mills, J., Bonner, A., and Francis, K. (2006b). The development of constructivist grounded theory. *Int. J. Qualitat. Methods* 5, 25–35. doi: 10.1177/160940690600500103

Miyahara, K., Niikawa, T., Hamada, H. T., and Nishida, S. (2020). Developing a short-term phenomenological training program: a report of methodological lessons. *New Ideas Psychol.* 58:100780. doi: 10.1016/j.newideapsych.2020.100780

Mogensen, J., and Overgaard, M. (2018). Reorganization of the connectivity between elementary functions as a common mechanism of phenomenal consciousness and working memory: from functions to strategies. *Philos. Trans. R. Soc. Lond. Ser. B Biol. Sci.* 373:20170346. doi: 10.1098/rstb.2017.0346

Moran, T. P. (2016). Anxiety and working memory capacity: a meta-analysis and narrative review. *Psychol. Bull.* 142, 831–864. doi: 10.1037/bul0000051

Morrison, H., McBriar, S., Powell, H., Proudfoot, J., Stanley, S., Fitzgerald, D., et al. (2019). What is a psychological task? The operational pliability of "Task" in psychological laboratory experimentation. *Engaging Sci. Technol. Soc.* 5, 61–85. doi: 10.17351/ests2019.274

Nelson, J. (2017). Using conceptual depth criteria: addressing the challenge of reaching saturation in qualitative research. *Qualitat. Res.* 17, 554–570. doi: 10.1177/1468794116679873

Nelson, J. K., Reuter-Lorenz, P. A., Sylvester, C. C., Jonides, J., and Smith, E. E. (2003). Dissociable neural mechanisms underlying response-based and familiarity-based conflict in working memory. *Proc. Natl. Acad. Sci. U.S.A.* 100, 11171–11175. doi: 10.1073/pnas.1334125100

Oblak, A. (2020). Visual representation in the wild: empirical phenomenological investigation of visual-spatial working memory in a naturalistic setting. *Constructivist Foundations* 15, 238–250.

Oblak, A. (2021b). The hell of being who one ordinarily is: is it possible to construct stable phenomenological traits of mood disorders? *Constructivist Foundations* 16, 227–229.

Oblak, A. (2021a). Participatory sense-making as consensual validation of phenomenal data. *Interdisciplinary Description Complex Syst.* 19, 470–492. doi: 10.7906/indecs.19.4.2

Oblak, A., Boyadzhieva, A., and Bon, J. (2021). Phenomenological properties of perceptual presence: a constructivist grounded theory approach. *Constructivist Foundations* 16, 295–308.

Orne, M. (1962). On the social psychology of the psychological experiment: with particular reference to demand characteristics and their implications. *Am. Psychol.* 17, 776–783. doi: 10.1037/h0043424

Peirce, J. W. (2007). PsychoPy - Psychophysics software in Python. *J. Neurosci. Methods* 162, 8–13. doi: 10.1016/j.jneumeth.2006.11.017

Petitmengin, C. (2006). Describing one's subjective experience in the second person: an interview method for the science of consciousness. *Phenomenol. Cogn. Sci.* 5, 229–269. doi: 10.1007/s11097-006-9022-2

Petitmengin, C. (2017). Enaction as a lived experience: towards a radical neurophenomenology. *Constructivist Foundations* 12, 139–147.

Petitmengin, C., Bitbol, M., Nissou, J., Pachoud, B., Curallucci, H., Cermolacce, M., et al. (2009). Listening from within. *J. Conscious. Stud.* 16, 252–284.

Petitmengin, C., Navarro, V., and Le Van Queyen, M. (2007). Anticipating seizure: pre-reflective experience at the center of neuro-phenomenology. *Conscious. Cogn.* 16, 746–764. doi: 10.1016/j.concog.2007.05.006

Petitmengin, C., Remillieux, A., and Valenzuela-Moguillansky, C. (2019). Discovering the structures of lived experience. *Phenomenol. Cogn. Sci.* 18, 691–730. doi: 10.1007/s11097-018-9597-4

Ratcliffe, M. (2008). *International Perspectives in Philosophy and Psychiatry. Feelings of Being: Phenomenology, Psychiatry and the Sense of Reality*. Oxford: Oxford University Press.

Repovš, G., and Baddeley, A. D. (2006). The multi-component model of working memory: explorations in experimental cognitive psychology. *Neuroscience* 139, 5–21. doi: 10.1016/j.neuroscience.2005.12.061

Rose, E. J., and Ebmeier, K. P. (2006). Pattern of impaired working memory during major depression. *J. Affect. Disord.* 2, 149–161. doi: 10.1016/j.jad.2005.11.003

Sanfratello, L., Caprihan, A., Stephen, J. M., Knoefel, J. E., Adair, J. C., Qualls, C., et al. (2014). Same task, different strategies: how brain networks can be influenced by memory strategy. *Hum. Brain Mapp.* 35, 5127–5140. doi: 10.1002/hbm.22538

Saunders, B., Sim, J., Kingstone, T., Baker, S., Waterfield, J., Bartlam, B., et al. (2018). Saturation in qualitative research: exploring its conceptualization and operationalization. *Quality Quantity* 52, 1893–1907. doi: 10.1007/s11135-017-0574-8

Schwartenbeck, P., Fitzgerald, T., Raymond, J. D., and Friston, K. J. (2013). Exploration, novelty, surprise, and free energy minimization. *Front. Psychol.* 4:710. doi: 10.3389/fpsyg.2013.00710

Schwartzman, D. J., Oblak, A., Rothen, N., Bor, D., and Seth, A. K. (2020). Extensive phenomenological overlap between induced and naturally-occurring synaesthetic experiences. *BioRxiv [Preprint]* 228692. doi: 10.1101/2020.08.03.228692

Seghier, M. L., and Price, C. J. (2018). Interpreting and utilising intersubject variability in brain function. *Trends Cogn. Sci.* 22, 517–530. doi: 10.1016/j.tics.2018.03.003

Seth, A. K. (2014). A predictive processing theory of sensorimotor contingencies: explaining the puzzle of perceptual presence and its absence in synesthesia. *Cogn. Neurosci.* 5, 97–118. doi: 10.1080/17588928.2013.877880

Seth, A. K., Suzuki, K., and Critchley, H. D. (2012). An interoceptive predictive coding model of conscious presence. *Front. Psychol.* 2:395. doi: 10.3389/fpsyg.2011.00395

Silverstein, M. S., Demmin, D., and Škodlar, B. (2017). Space and objects: on the phenomenology and cognitive neuroscience of anomalous perception in schizophrenia (Ancillary Article to EAWE Domain 1). *Psychopathology* 50, 60–67. doi: 10.1159/000452493

Soto, D., Mäntylä, T., and Silvanto, J. (2011). Working memory without consciousness. *Curr. Biol.* 21, R912–R913. doi: 10.1016/j.cub.2011.09.049

Soto, D., and Silvanto, J. (2014). Reappraising the relationship between working memory and conscious awareness. *Trends Cogn. Sci.* 18, 520–525. doi: 10.1016/j.tics.2014.06.005

Speer, N. K., Jacoby, L. L., and Braver, T. S. (2003). Strategy-dependent changes in memory: effects on behavior and brain activity. *Cogn. Affect. Behav. Neurosci.* 3, 155–167. doi: 10.3758/CABN.3.3.155

Starc, M., Anticevic, A., and Repovš, G. (2017). Fine-grained versus categorical: pupil size differentiates between strategies for spatial working memory performance. *Psychophysiology* 54, 724–735. doi: 10.1111/psyp.12828

Stone, C. (2019). A defense and definition of comstruct validity in psychology. *Philos. Sci.* 86, 1250–1261. doi: 10.1086/705567

Thompson, E. (2007). *Mind in Life: Biology, Phenomenology, and the Sciences of the Mind*. Cambridge: Harvard University Press.

Trigg, D. (2017). *Topophobia: A phenomenology of anxiety*. London: Bloomsbury.

Tsuchiya, N., and Koch, C. (2016). "The relationship between consciousness and top-down attention," in *The Neurology of Consciousness: Cognitive Neuroscience and Neuropathology*, eds S. Laureys, O. Gosseries, and G. Tononi (San Diego, CA: Academic Press), 71–91. doi: 10.1016/b978-0-12-800948-2.00005-4

Tsukiura, T., Mochizuki-Kawai, H., and Fujii, T. (2005). The effect of encoding strategies on medial temporal lobe activations during the recognition of words: an event-related fMRI study. *Neuroimage* 25, 452–461. doi: 10.1016/j.neuroimage.2005.01.003

Valenzuela-Moguillansky, C., and Demšar, E. (2022). Toward a science of experience: outlining some challenges and future directions. *Adapt. Behavior.* doi: 10.1177/10597123221084739

Valenzuela-Moguillansky, C., O'Regan, J. K., and Petitmengin, C. (2013). Exploring the subjective experience of the "rubber hand" illusion. *Front. Hum. Neurosci.* 7:659. doi: 10.3389/fnhum.2013.00659

Valenzuela-Moguillansky, C., and Vásquez-Rosati, A. (2019). An analysis procedure for the micro-phenomenological interview. *Constructivist Foundations* 14, 123–145.

Varela, F. J. (1996). Neurophenomenology: a methodological remedy for the hard problem. *J. Conscious. Stud.* 3, 330–349.

Varela, F. J., Tompson, E., and Rosch, E. (1991). *The Embodied Mind: Cognitive Science and Human Experience*. Cambridge: The MIT Press.

Wagemann, J. (2021). Exploring the structure of mental action in directed thought. *Philos. Psychol.* 35, 145–176. doi: 10.1080/09515089.2021.1951195

Weger, U., and Wagemann, J. (2015). The challenges and opportunities of first-person inquiry in experimental psychology. *New Ideas Psychol.* 36, 38–49. doi: 10.1016/j.newideapsych.2014.09.001

Yarkoni, T. (2022). The generalizability crisis. *Behav. Brain Sci.* 45:E1. doi: 10.1017/S0140525X20001685

Conflict of Interest: GR consults for and holds equity in Neumora Therapeutics.

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

frontiers | Frontiers in Psychology

MINI REVIEW
published: 02 June 2022
doi: 10.3389/fpsyg.2022.917139

Feeling Safe With Hypnosis: Eliciting Positive Feelings During a Special State of Consciousness

Barbara Schmidt[1,2]*

[1] *Institute for Psychosocial Medicine, Psychotherapy and Psychooncology, Jena University Hospital, Jena, Germany,*
[2] *Institute for Clinical Psychology, University of Jena, Jena, Germany*

Our state of consciousness is crucial for our ability to follow suggestions. Suggestions in turn are a powerful tool to induce positive emotional states. In my research, I suggest positive feelings of safety during hypnosis. This is a positive emotional state of low arousal and low anxiety. Both arousal and anxiety affect our decision-making. However, when we feel safe due to hypnotic suggestions of safety, we do not act riskier. Instead, EEG brain activity shows that monetary rewards get less important and delayed rewards are less devalued compared to immediate rewards when we feel safe. These results open promising perspectives for the use of hypnosis to reduce impulsive behavior, for example, in substance abuse. Therapeutic suggestions of safety even work in highly stressful environments like the intensive care unit. I showed that patients tolerate non-invasive ventilation much better when they get the suggestion to feel safe. The effects of positive therapeutic suggestions delivered during hypnosis even persist over time. Post-hypnotic suggestions are associations between a certain emotional state and a trigger that elicits this emotional state after hypnosis is over. I showed that post-hypnotic suggestions of safety are effective weeks after the therapeutic session. Therefore, I present a therapeutic technique that uses a special state of consciousness, hypnosis, to induce positive emotional states. The effects of this technique are very strong and long lasting. My goal is to provide scientific evidence for the use of hypnotherapeutic techniques to increase the number of people who apply and profit from them.

Keywords: hypnosis, anxiety, arousal, decision-making, intensive care unit, delay discounting, risk behavior, safety

OPEN ACCESS

Edited by:
Antonino Raffone,
Sapienza University of Rome, Italy

Reviewed by:
Lucie Bréchet,
University of Geneva, Switzerland

****Correspondence:***
Barbara Schmidt
Barbara.Schmidt@med.uni-jena.de
orcid.org/0000-0002-6828-5546

Specialty section:
This article was submitted to Consciousness Research, a section of the journal Frontiers in Psychology

Received: *10 April 2022*
Accepted: *16 May 2022*
Published: *02 June 2022*

Citation:
Schmidt B (2022) Feeling Safe With Hypnosis: Eliciting Positive Feelings During a Special State of Consciousness. Front. Psychol. 13:917139. doi: 10.3389/fpsyg.2022.917139

HYPNOSIS AS A SPECIAL STATE OF CONSCIOUSNESS

Different states of consciousness are a normal part of our life. When we sleep, we are in a certain state of consciousness. When we are in the middle of doing something that we really like, we can also get into a very special state of consciousness called trance or flow. A trance or flow state of consciousness enables effortless performance and is associated with positive feelings. Musicians often perceive this state of consciousness as an optimal balance between their skills and the challenge they are facing, an experience which is related to high performance and incompatible with stage anxiety (Cohen and Bodner, 2019). Professional athletes experiencing flow report to have a clear idea of their goals and get unambiguous feedback, which in turn predicts high satisfaction with life (Habe et al., 2019). High performance is often accompanied by a trance state where it is possible to show your optimal performance, because you are totally absorbed by the moment and nothing else matters. It is a state of very intense focus on the one thing that matters while ignoring all other irrelevant stimuli. Some even use potentially distracting stimuli to get deeper into their optimal state of performance, a technique described by one of the most important pioneers in hypnosis, Milton Erickson (1959).

When we elicit this trance state *via* a hypnosis induction, we call it hypnosis. The current definition of hypnosis was stated by the APA Division 30 in the year 2015. The division defines hypnosis as follows: "A state of consciousness involving focused attention and reduced peripheral awareness characterized by an enhanced capacity for response to suggestion" (Elkins et al., 2015). Suggestions are contents that the hypnotist says. An example for a suggestion is: "The longer you look at this point in front of you, the heavier get your eyelids so that you want to close them." The hypnotized person can then accept this suggestion and close his or her eyes. Following a suggestion feels like an automatic process instead of a conscious decision. It feels natural to follow the suggestion and close your eyes. The ability to follow suggestions is dependent on the rapport between the hypnotist and the hypnotized person, describing a positive therapeutic relationship of trust and responsibility. Suggestibility is usually measured with the Harvard Group Scale of Hypnotic Susceptibility (HGSHS; Shor and Orne, 1963). This is a group test consisting of a hypnosis induction and 12 suggestions. Participants indicate on a questionnaire if they followed the suggestions, resulting in a score from 0 for very low suggestibility to 12 for very high suggestibility.

Hypnosis as a special state of consciousness has long been fascinating scientists. Ivan Pavlov reported that some of his dogs were in a trance-like state after his experimental procedures of classical conditioning (Pavlov and Petrova, 1934). Hans Jürgen Eysenck studied the differences between primary and secondary suggestibility, which describes the ability to follow suggestions (Eysenck and Furneaux, 1945). After the invention of the electroencephalogram (EEG) in Jena (Berger, 1929), the famous Berger effect was investigated with hypnosis. In Berger's original experimental setup, the participant was opening and closing his eyes while Berger measured the participant's EEG alpha activity. Alpha activity was only visible in the EEG signal when the participant's eyes were closed and blocked when eyes were open. In the hypnosis study, the participant was suggested during hypnosis that he was blind or that he can see. When the participant was suggested to be blind, EEG alpha activity was visible, even though the eyes were open (Loomis et al., 1936). This observation shows that sensory processes can be altered with hypnotic suggestions.

BLOCKING SENSORY PERCEPTION UNDER HYPNOSIS

In three separate EEG studies, I blocked sensory perception of visual, pain, and auditory stimuli using hypnotic suggestions of a wooden board in front of participants' eyes, a cooling and numbing glove on participants' hands and earplugs in participants' ears, respectively (**Figure 1**; Schmidt et al., 2017b; Franz et al., 2020, 2021). We used a sensory paradigm that is very common in EEG research with frequent and rare stimuli, called oddball paradigm. In the visual study, participants had to count the rare stimuli on the screen, which were colored squares in my study. Brain responses usually show a very clear response to the rare to be attended stimuli in the so called P3 response. That is a positive voltage change about 300 milliseconds after the stimulus was presented. I had three groups of participants in this study, selected according to their HGSHS suggestibility scores (Shor and Orne, 1963). I had 20 low suggestible, 20 middle suggestible and 20 high suggestible participants. All participants played the oddball task twice in counterbalanced order. Once with hypnosis and the suggestion that a wooden board is blocking their vision (**Figure 1A**), once in a control condition without hypnosis. In both conditions, participants saw the stimuli of the oddball task on the screen, presented one at a time, and counted the rare colored squares. While participants sat in the EEG chamber, I checked that their eyes were open all the time. The results show that participants were not able to correctly count the rare stimuli in the hypnosis condition, while they showed almost perfect counting performance in the control condition. Participants' brain responses showed that the visual stimuli were still perceived in the hypnosis condition, indicated by early event-related EEG components. But the target P3 component was massively reduced in the hypnosis condition. The smaller the P3 amplitude was, the more reduced was participants' counting performance and the more vivid was their experience of the wooden board in front of their eyes in the hypnosis condition. The effect was strongest for high suggestible participants. The results show the neuronal dissociation between perceiving the visual objects on the screen and attending them in order to count them, reflected in early and late event-related EEG signals. In a subsequent analysis, we found that this effect was driven by a top-down modulation, reflected in reduced directed information flow from parietal attentional to frontal executive sources during processing of target stimuli (Franz et al., 2021).

We obtained similar results in the two subsequent EEG studies where I blocked pain and auditory processing with hypnotic suggestions. In the pain study, I blocked participants' pain processing *via* the suggestion of a glove that keeps the stimulated hand from feeling pain (**Figure 1B**) similar to the suggestion by Rainville et al. (1997). In the auditory study, I blocked participants' auditory processing *via* the suggestion of earplugs (Franz et al., 2020). In these studies, I included additional control conditions of attention distraction and simulation of hypnosis. In the auditory study, we used an auditory oddball paradigm where participants are presented frequent and rare sounds. Participants had to press a button when they heard the rare target sound. In participants' EEG, we focused again on the P3 component to the rare target sounds. We found a significant reduction in P3 amplitudes in the hypnosis condition compared to the control condition. Participants also pressed the button to the target sound significantly less often and perceived the sounds as less loud in the hypnosis condition compared to the control

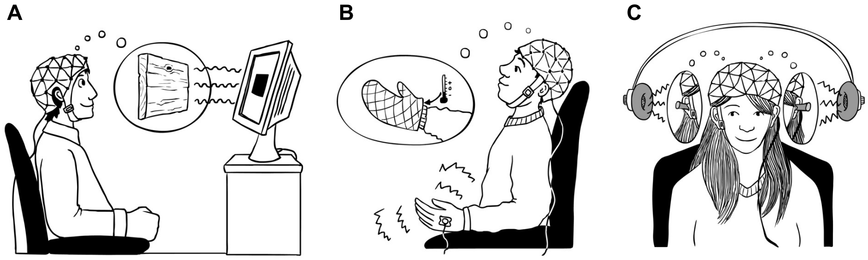


FIGURE 1 | Illustration of the three EEG studies using hypnotic suggestions to block sensory processing. **(A)** Participants had to count rare visual stimuli on the screen. In the hypnosis condition, they were suggested that a wooden board is blocking their vision on the screen, so they cannot see the visual stimuli. **(B)** Participants received electrical pain stimuli on their hand. In the hypnosis condition, they were suggested that a cooling and numbing glove is covering their hand, so they cannot feel the stimuli. **(C)** Participants had to press a button when they heard a rare sound. In the hypnosis condition, they were suggested that earplugs keep them from hearing the sounds.

condition (Franz et al., 2020). Taken together, the results of the three sensory blockade studies reveal that hypnotic suggestions are a powerful tool to modify sensory processes in the brain, especially processes that are associated with attention control and stimulus evaluation like the P3 amplitude.

HYPNOSIS, AROUSAL, AND DECISION-MAKING

The state of hypnosis is not only characterized by an enhanced ability to follow suggestions, but also by low arousal. When you are in hypnosis, you are relaxed. Therefore, hypnosis was a valuable tool for the development of systematic desensitization as described by Wolpe et al. (1973). In his description of the first standardized technique in psychotherapy, Wolpe uses hypnosis to relax the patient before the patient imagines the objects or situations that he or she is afraid of. The imagination of fear stimuli can be as efficient as real fear stimuli, as revealed by a recent fear conditioning study (Mueller et al., 2019). In this study, participants imagined stepping into a thumbtack when a certain visual stimulus appeared on the screen. Participants developed a conditioned fear response like in previous fear conditioning studies that used real instead of imagined stimuli. The study by Mueller et al. (2019) provides further evidence for the effectiveness of imagination. To reduce fear responses, participants can use hypnosis to get relaxed and then imagine the previously fear-eliciting stimulus. As Wolpe et al. (1973) report, this is a very effective method to reduce anxiety.

Reducing participants' arousal typically affects their decision-making behavior. I showed that participants who have generally lower arousal, indicated by a low resting heart rate, acted riskier in a risk game than participants with higher arousal (Schmidt et al., 2013). When participants' state arousal was increased after riding the bike for 10 min on a bicycle home-trainer, they tended to be less risky (Schmidt et al., 2013). Being aroused often goes along with being anxious. I found that more anxious participants acted less risky in a risk game and showed higher frontal midline theta power than less anxious participants (Schmidt et al., 2018). When participants wore a bike helmet, they showed lower frontal midline theta power, but did not generally act riskier in a risk game compared to participants without bike helmet (Schmidt et al., 2019). These studies show that lower arousal is associated with less anxiety and riskier behavior. Reducing arousal with hypnosis might therefore also reduce anxiety and affect decision-making.

FEELING SAFE WITH HYPNOSIS

One of the most prominent techniques in hypnotherapy is to suggest participants that they are at a safe place. The need for safety was stated as one of our most basic needs by Abraham Maslow (1943). In his seminal publication that resulted in his famous pyramid showing the hierarchy of needs, he states: "Practically everything looks less important than safety (even sometimes the physiological needs which being satisfied, are now underestimated). A man, in this state, if it is extreme enough and chronic enough, may be characterized as living almost for safety alone" (Maslow, 1943). The safe place hypnosis technique uses the suggestion that the hypnotized person is at his or her personal safe place (Arntz, 2011). The hypnotized person is free to choose his or her own imaginations, the hypnotist only offers suggestions. One example would be: "Be curious what you can see, hear and smell at your safe place. Feel the place in your body. Where is this feeling most intense? Focus on this part of your body and make the feeling grow even stronger. It radiates through your whole body" (Schmidt et al., 2020).

I developed a standardized safe place suggestion and measured brain responses and behavior of participants in a risk game. For this study, I only invited highly suggestible participants that were again pre-tested with the HGSHS (Shor and Orne, 1963). Participants played a risk game twice in two conditions while recording their EEG brain responses (Schmidt et al., 2020). In the hypnosis and safety condition, I hypnotized participants and suggested them to be at their own safe place. Then, they played the risk game. In the control condition, participants played the risk game without hypnosis. To understand the results of this study, it is important to know that monetary rewards elicit a P3 response. Higher monetary rewards are reflected in higher P3 responses (Begleiter et al., 1983). The same is true for other rewarding stimuli. For example, smokers respond with a strong

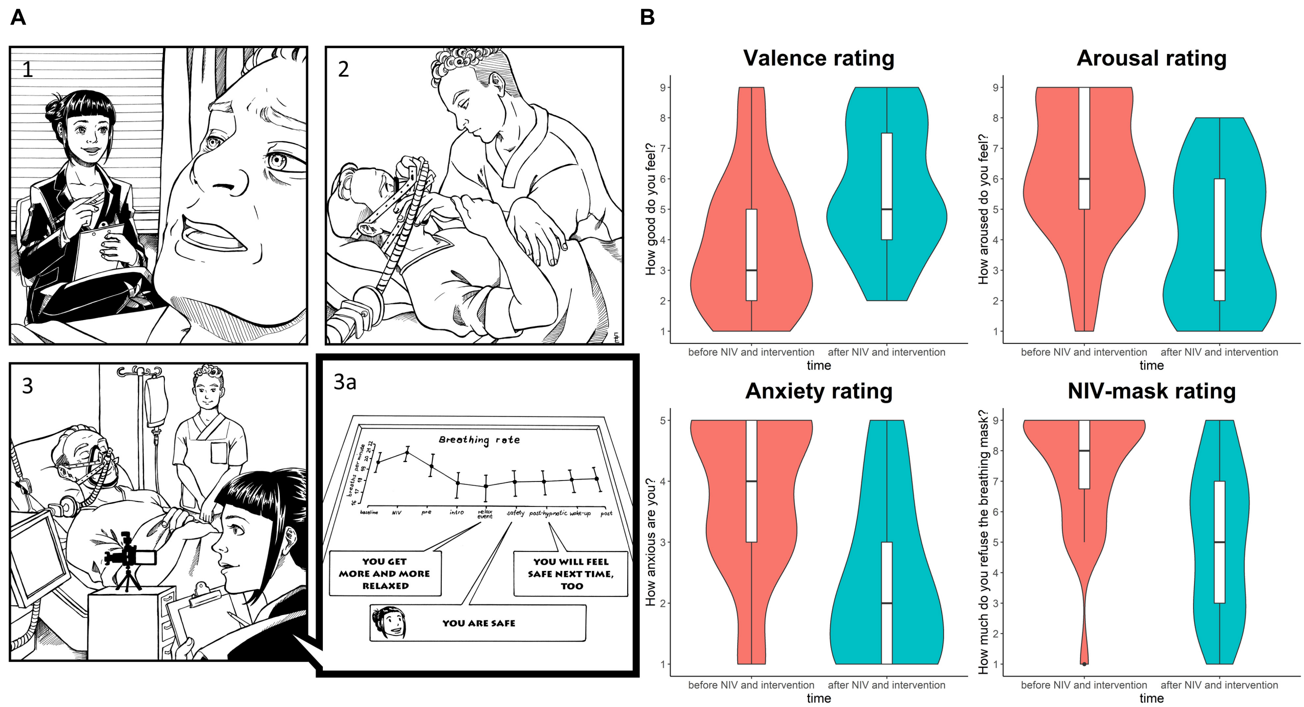


FIGURE 2 | (A) Illustration of the procedure of the intensive care study. The psychologist asked the patient about his current emotional state. Then, a member of the medical staff puts on the NIV breathing mask. During the ventilation session, the psychologist accompanies the patient with therapeutic suggestions of safety. The vital sign monitor shows that breathing rate was reduced during safe place suggestions. **(B)** Results of ratings before and after NIV sessions accompanied with safe place suggestions showing better mood, less arousal and anxiety, and a more positive evaluation of the NIV mask after the intervention.

P3 amplitude to images related to smoking compared to other images. In my study, I used monetary rewards as incentives. My results show that participants showed significantly lower P3 amplitudes to all monetary rewards, large or small, when they felt safe in the hypnosis condition compared to the control condition. Excitingly, similarly reduced P3 amplitudes were reported in smokers who no longer smoke (Littel and Franken, 2007). While smokers in the study had to be abstinent for a long time to stop having strong P3 responses to the smoke pictures, in my study only one hypnosis session with safety suggestion was sufficient. Imagining a safe place may be a way to stop having strong responses to reward stimuli. This could aid in the treatment of addictive disorders. If people stop reacting strongly to stimulants to which they are addicted, they will find it easier to give up consuming these stimulants. Importantly, risk behavior did not change when participants felt safe during hypnosis. Therefore, there is no contraindication to use suggestions of safety in patients suffering from addiction.

In a second paradigm, I investigated another phenomenon, the devaluation of future rewards, also called delay discounting (Schmidt and Holroyd, 2021). You are probably familiar with the famous marshmallow task by Mischel et al. (1988). A preschool-aged child is presented with the task of either eating one marshmallow now or waiting to get two marshmallows. Children who were able to wait for the second marshmallow showed better social and academic performance later (Mischel et al., 1988). Thus, the ability to wait for later rewards is desirable. In this context, children's decisions to wait for rewards also depends on how much they trust their environment (Mahrer, 1956; Kidd et al., 2013). I therefore hypothesized that individuals who feel safe would be more willing to wait. In the delayed gratification game that I used in my EEG study (Schmidt and Holroyd, 2021), participants could win immediate monetary rewards and rewards that were paid 6 months later. The EEG brain activity shows more positive deflections after an immediate reward than after a delayed reward. The difference between the deflection of the EEG signal to immediate and delayed rewards is called reward positivity. If this difference is large, our brain makes a strong distinction between immediate and delayed rewards, and it is difficult to wait. However, if this difference is small, we will find it easier to wait. I found almost no difference between EEG brain responses to immediate and delayed rewards in participants who felt safe. In contrast, they showed strong EEG differences in the control condition. When I compare the results of this study with the results of an earlier study using the same paradigm (Schmidt et al., 2017a), it becomes even clearer how exciting these findings are. In the earlier study, I compared two groups of participants. One group was low impulsive and high self-controlled, and the other group was high impulsive and low self-controlled. Participants in the low impulsivity group showed a comparably small difference between immediate and delayed rewards as did the participants in my current study (Schmidt and Holroyd, 2021) when they felt safe during hypnosis. And this was after a single hypnosis session with the suggestion of being at a safe place. It might therefore be possible to make participants less impulsive and more controlled by suggesting them to be at a safe place. This, in turn, could make it easier for them to wait for future rewards. Forgoing immediate rewards in favor of future rewards requires a high degree of self-control and confidence that waiting will pay off. This skill is essential in coping with individual problems such as substance addiction and obesity, as well as global problems such as climate change and a pandemic.

After these encouraging results, I decided to link the feeling of safety to a post-hypnotic trigger (Böhmer and Schmidt, 2022). I wanted to investigate whether this could trigger the feeling of being in a safe place in the long term without the need for another therapeutic session. In order to establish a safety trigger, I used a white sheet of paper on which participants under hypnosis wrote the letter S for safety. I suggested the participants that the feeling of safety was stored in the piece of paper. Every time they looked at this piece of paper, folded it up and put it into their pocket, the feeling of safety should automatically reappear. In the EEG study, everything was similar to what I described earlier (Schmidt et al., 2020; Schmidt and Holroyd, 2021), except that I ended the hypnotic state before playing the risk game in the safety condition. Participants were then given the S paper or a neutral paper with the letter K for control on it. They again played the risk game twice, once with the post-hypnotic suggestion of safety and once with the neutral paper as the control condition. Again, participants' risk behavior did not differ in both conditions. We analyzed the difference between the EEG brain responses to high and low monetary rewards. A large difference here is an indicator of strong reward sensitivity. The brain then responds strongly to rewards. We found that participants' EEG brain responses in the safety condition made almost no difference between high and low monetary rewards, while in the control condition they made a very strong distinction between high and low rewards. The reduced responsiveness for reward stimuli indicates a state of satisfaction and satiation. Such a state is very helpful in treating individuals who are otherwise very responsive to reward stimuli such as individuals suffering from substance abuse. We also asked participants weeks after the initial experimental session if the post-hypnotic safety trigger still worked. Participants indicated that the S piece of paper still elicited a feeling of safety, showing the long-lasting effects of post-hypnotic suggestions.

FEELING SAFE IN THE INTENSIVE CARE UNIT WITH HYPNOTIC SUGGESTIONS OF SAFETY

To prove the effectiveness of a therapeutic suggestion, it is important to show that it works in naturally occurring situations. I therefore did a study in which we used the safe place method inside the intensive care unit (Schmidt et al., 2021). In this study, my master's student Jana Schneider accompanied patients who have strong fear of non-invasive ventilation. Non-invasive ventilation can cause feelings of suffocation when trying to breathe against the machine, even though ventilation is intended to ensure that the patient's body is optimally supplied with oxygen. Regular ventilation sessions in the intensive care unit last about 15 min. We accompanied the patients during one of those ventilation sessions. The ventilation and the suggestion of the safe place thus took place simultaneously. We included only patients who were awake and able to provide information about their current state. Before and after the ventilation session, we asked patients how anxious they were, how aroused they were, and how well they generally felt. During ventilation, we recorded the patient's vital signs monitor. This allowed us to analyze exactly how the patients' bodies responded to the safe place suggestion. We found that the respiratory rate was reduced as a sign of relaxation during the safe place suggestion and the heart rate also calmed down. After the intervention, patients reported feeling less anxious, less aroused, and generally feeling better. They also rated the breathing mask itself as less negative. **Figure 2** shows the procedure and results of this study.

CONCLUSION

In my studies I showed that the suggestion of a safe place is very effective both in the EEG laboratory and in the intensive care unit during challenging medical procedures. The effects were particularly large in the intensive care unit, where we assume a naturally occurring trance state that contributes to the effectiveness of the suggestions. From my EEG studies, I can draw conclusion about the effect of safety suggestion under hypnosis and as a post-hypnotic suggestion. Immediately after hypnosis, subjects felt safer than with post-hypnotic suggestion, with the effectiveness of post-hypnotic suggestion lasting for weeks. In summary, the studies confirm the high efficacy and good applicability of the safe place therapeutic technique. It is my sincere wish that through my research I will contribute to the even more widespread use of this technique and help even more people to turn fear into safety.

AUTHOR CONTRIBUTIONS

BS wrote the manuscript, contributed to the article, and approved the submitted version.

REFERENCES

Arntz, A. (2011). Imagery rescripting for personality disorders. *Cogn. Behav. Pract.* 18, 466–481. doi: 10.1016/j.cbpra.2011.04.006

Begleiter, H., Porjesz, B., Chou, C. L., and Aunon, J. I. (1983). P3 and stimulus incentive value. *Psychophysiology* 20, 95–101. doi: 10.1111/j.1469-8986.1983.tb00909.x

Berger, H. (1929). Über das Elektrenkephalogramm des Menschen. *Archiv für Psychiatrie und Nervenkrankheiten* 87, 527–570. doi: 10.1007/BF01797193

Böhmer, J., and Schmidt, B. (2022). Safety on demand: Post-hypnotic suggestions of safety reduce neural signals of reward sensitivity with long-term beneficial effects. *Psychophysiology* e14015. doi: 10.1111/psyp.14015 [Epub ahead of print].

Cohen, S., and Bodner, E. (2019). The relationship between flow and music performance anxiety amongst professional classical orchestral musicians. *Psychol. Music* 47, 420–435. doi: 10.1177/0305735618754689

Elkins, G. R., Barabasz, A. F., Council, J. R., and Spiegel, D. (2015). Advancing research and practice: The revised APA Division 30 definition of hypnosis. *Am. J. Clin. Hypnosis* 57, 378–385. doi: 10.1080/00029157.2015.1011465

Erickson, M. H. (1959). Further clinical techniques of hypnosis: Utilization techniques. *Am. J. Clin. Hypnosis* 2, 3–21. doi: 10.1080/00029157.1959.10401792

Eysenck, H. J., and Furneaux, W. D. (1945). Primary and secondary suggestibility: An experimental and statistical study. *J. Exp. Psychol.* 35, 485–503. doi: 10.1037/h0054976

Franz, M., Schmidt, B., Hecht, H., Naumann, E., and Miltner, W. H. R. (2020). Suggested deafness during hypnosis and simulation of hypnosis compared to a distraction and control condition: A study on subjective experience and cortical brain responses. *PLoS One* 15:e0240832. doi: 10.1371/journal.pone.0240832

Franz, M., Schmidt, B., Hecht, H., Naumann, E., and Miltner, W. H. R. (2021). Suggested visual blockade during hypnosis: Top-down modulation of stimulus processing in a visual oddball task. *PLoS One* 16:e0257380. doi: 10.1371/journal.pone.0257380

Habe, K., Biasutti, M., and Kajtna, T. (2019). Flow and satisfaction with life in elite musicians and top athletes. *Front. Psychol.* 10:698. doi: 10.3389/fpsyg.2019.00698

Kidd, C., Palmeri, H., and Aslin, R. N. (2013). Rational snacking: Young children's decision-making on the marshmallow task is moderated by beliefs about environmental reliability. *Cognition* 126, 109–114. doi: 10.1016/j.cognition.2012.08.004

Littel, M., and Franken, I. H. A. (2007). The effects of prolonged abstinence on the processing of smoking cues: an ERP study among smokers, ex-smokers and never-smokers. *J. Psychopharmacol.* 21, 873–882. doi: 10.1177/0269881107078494

Loomis, A. L., Harvey, E. N., and Hobart, G. (1936). Brain potentials during hypnosis. *Science* 83, 239–241. doi: 10.1126/science.83.2149.239

Mahrer, A. R. (1956). The role of expectancy in delayed reinforcement. *J. Exp. Psychol.* 52, 101–106. doi: 10.1037/h0040837

Maslow, A. H. (1943). A theory of human motivation. *Psychol. Rev.* 50, 370–396. doi: 10.1037/h0054346

Mischel, W., Shoda, Y., and Peake, P. K. (1988). The nature of adolescent competencies predicted by preschool delay of gratification. *J. Personal. Soc. Psychol.* 54, 687–696. doi: 10.1037/0022-3514.54.4.687

Mueller, E. M., Sperl, M. F. J., and Panitz, C. (2019). Aversive imagery causes de novo fear conditioning. *Psychol. Sci.* 30, 1001–1015. doi: 10.1177/0956797619842261

Pavlov, I. P., and Petrova, M. K. (1934). A contribution to the physiology of the hypnotic state of dogs. *J. Personal.* 2, 189–200. doi: 10.1111/j.1467-6494.1934.tb02095.x

Rainville, P., Duncan, G. H., Price, D. D., Carrier, B., and Bushnell, M. C. (1997). Pain affect encoded in human anterior cingulate but not somatosensory cortex. *Science* 277, 968–971. doi: 10.1126/science.277.5328.968

Schmidt, B., Hecht, H., Naumann, E., and Miltner, W. H. R. (2017b). The power of mind: Blocking visual perception by hypnosis. *Sci. Rep.* 7:4889. doi: 10.1038/s41598-017-05195-2

Schmidt, B., Hoffmann, E., and Rasch, B. (2020). Feel safe and money is less important! Hypnotic suggestions of safety decrease brain responses to monetary rewards in a risk game. *Cereb. Cortex Commun.* 1, 1–9. doi: 10.1093/texcom/tgaa050

Schmidt, B., and Holroyd, C. B. (2021). Hypnotic suggestions of safety reduce neuronal signals of delay discounting. *Sci. Rep.* 11:2706. doi: 10.1038/s41598-021-81572-2

Schmidt, B., Holroyd, C. B., Debener, S., and Hewig, J. (2017a). I can't wait! Neural reward signals in impulsive individuals exaggerate the difference between immediate and future rewards. *Psychophysiology* 54, 409–415. doi: 10.1111/psyp.12796

Schmidt, B., Kanis, H., Holroyd, C. B., Miltner, W. H. R., and Hewig, J. (2018). Anxious gambling: Anxiety is associated with higher frontal midline theta predicting less risky decisions. *Psychophysiology* 55:e13210. doi: 10.1111/psyp.13210

Schmidt, B., Kessler, L., Holroyd, C. B., and Miltner, W. H. R. (2019). Wearing a bike helmet leads to less cognitive control, revealed by lower frontal midline theta power and risk indifference. *Psychophysiology* 56:e13458. doi: 10.1111/psyp.13458

Schmidt, B., Mussel, P., and Hewig, J. (2013). I'm too calm – let's take a risk! On the impact of state and trait arousal on risk taking. *Psychophysiology* 50, 498–503. doi: 10.1111/psyp.12032

Schmidt, B., Schneider, J., Deffner, T., and Rosendahl, J. (2021). Hypnotic suggestions of safety improve well-being in non-invasively ventilated patients in the intensive care unit. *Intens. Care Med.* 47, 485–486. doi: 10.1007/s00134-021-06364-8

Shor, R. E., and Orne, E. C. (1963). Norms on the Harvard Group Scale of Hypnotic Suggestibility, Form A. *Int. J. Clin. Exp. Hypnosis* 11, 39–47. doi: 10.1080/00207146308409226

Wolpe, J., Brady, J. P., Serber, M., Agras, W. S., and Liberman, R. P. (1973). The current status of systematic desensitization. *Am. J. Psychiatry* 130, 961–965. doi: 10.1176/ajp.130.9.961

Conflict of Interest: The author declares that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's Note: All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

frontiers | Frontiers in Psychology

PERSPECTIVE
published: 22 June 2022
doi: 10.3389/fpsyg.2022.855450

Personal Memories and Bodily-Cues Influence Our Sense of Self

Lucie Bréchet[1,2]*

[1] *Department of Basic Neurosciences, University of Geneva, Geneva, Switzerland,* [2] *Department of Neurology, Harvard Medical School, Boston, MA, United States*

How do our bodies influence who we are? Recent research in cognitive neuroscience has examined consciousness associated with the self and related multisensory processing of bodily signals, the so-called bodily self-consciousness. A parallel line of research has highlighted the concept of the autobiographical self and the associated autonoetic consciousness, which enables us to mentally travel in time. The subjective re-experiencing of past episodes is described as re-living them from within or outside one's body. In this brief perspective, I aim to explore the underlying characteristics of self-consciousness and its relation to bodily signals and episodic memory. I will outline some recent behavioral and neuroimaging evidence indicating that bodily cues play a fundamental role in autobiographical memory. Finally, I will discuss these emerging concepts regarding the current understanding of bodily-self, autobiographical-self, their links to self-consciousness, and suggest directions for future research.

Keywords: bodily-self, autobiographical-self, out-of-body experiences, view-points, autonoetic consciousness, fMRI, VR

OPEN ACCESS

Edited by:
Antonino Raffone,
Sapienza University of Rome, Italy

Reviewed by:
Jane Elizabeth Aspell,
Anglia Ruskin University,
United Kingdom

Correspondence:
Lucie Bréchet
lucie.brechet@unige.ch

Specialty section:
This article was submitted to Consciousness Research, a section of the journal Frontiers in Psychology

Received: *15 January 2022*
Accepted: *02 June 2022*
Published: *22 June 2022*

Citation:
Bréchet L (2022) Personal Memories and Bodily-Cues Influence Our Sense of Self. Front. Psychol. 13:855450. doi: 10.3389/fpsyg.2022.855450

INTRODUCTION

The conscious experience of self-related events is perceived as embodied, i.e., our physical body is a central object in the world. We naturally experience and make sense of the world from an inside viewpoint of our physical body (Mach, 1897), yet we can retrieve memories from both within the body (i.e., own eyes/first-person perspective) or outside the body (i.e., observer/third-person perspective) (Nigro and Neisser, 1983; Rubin and Umanath, 2015). An essential aspect of consciousness is its link with a self, which is the subject of conscious experience. Self-consciousness originates at different levels, from a simple (i.e., the minimal level of unconscious experience) to a complex phenomenon (i.e., the core self and the autobiographical self) (Damasio, 1999). The core consciousness occurs when the brain continuously builds a mental representation of the minimal sense of self caused by an interaction with internal or external stimuli. This minimal sense of self is centered on integrating multisensory, interoceptive and exteroceptive bodily processing (e.g., vision, touch, proprioception, vestibular, and visceral signals) in the brain (Blanke et al., 2015; Park and Blanke, 2019). Manipulating the experience of one's own body experimentally has been challenging. As James (1890), the pioneering psychologist and philosopher, pointed out in the nineteenth century: *"The body is always there."* However, the recent advances in virtual reality technology enabled the experimental investigation of the global aspects of bodily self-consciousness, including the self-identification, self-location, and first-person perspective (Ehrsson, 2007; Lenggenhager et al., 2007). The bodily-self, which is the most fundamental aspect of self-consciousness, may influence higher cognitive aspects of self-representation, such as the autobiographical-self (Bergouignan et al., 2014; Bréchet et al., 2019, 2020; Tacikowski et al., 2020; Iriye and St. Jacques, 2021).

Conscious experiences, whether occurring from within or outside the body, are not always bound to here and now. What did you eat for breakfast yesterday? Where do you have to go today? When is your meeting tomorrow? The human mind can detach itself from the present moment and mentally travel to the past or imagine the future (Schacter et al., 2012). As Antonio Damasio (1999) defined, *the autobiographical self* represents a mental state derived from retrieving self-relevant memories. No scene captures the human ability to re-experience past events better than the well-known Marcel Proust's "madeleine moment." Struggling to remember details of his childhood, Marcel tastes a petite madeleine soaked in lime-blossom tea, "*and suddenly the memory revealed itself. The taste was that of the little piece of madeleine. . .had recalled nothing to my mind before I tasted it. And all from my cup of tea.*" Tulving (1985) associated the subjective possibility to mentally travel in time with autonoetic consciousness, i.e., the sense of self we experience when we subjectively re-experience an event and mentally travel in time. Mental time travel relies on episodic autobiographical memory, which allows humans to mentally detach themselves from a current self-location and consciously identify themselves at another particular place and time (Hassabis et al., 2007; Schacter et al., 2012; Zaman and Russell, 2021).

The relation between the self and episodic autobiographical memory lies at the core of our understanding of consciousness. One's autobiographical memories and sense of self are closely related. The self, i.e., the subjective feeling that defines us as unique human beings, is a critical component of consciousness (Damasio, 2003). Memory is necessary for everything we do and provides continuity from one moment to another. Memory creates our conscious sense of identity, which we construct based on self-relevant past events. As James (1890) suggested: *"I enter a friend's room and see on the wall a painting." At first, I have the strange, wondering consciousness, "surely I have seen that before," but when or how does not become clear. There only clings to the picture a sort of penumbra of familiarity—when suddenly I exclaim: "I have it, it is a copy of part of one of the Fra Angelico in the Florentine Academy—I recollect it there!"*

BODILY-SELF INFLUENCES THE AUTOBIOGRAPHICAL-SELF

Episodic autobiographical memory studies have been recently conducted using new approaches to control memory encoding and reflect accurate life-like testing outside the laboratory setting. For example, St Jacques and Schacter (2013), Nielson et al. (2015), and Vogel and Schwabe (2016) developed paradigms in which participants encoded real-life events while wearing a camera that automatically took photos. Still, these studies did not integrate the natural occurrence of participants' physical bodies during memory retrieval. Perceiving one's own physical body as part of a visual scene, such as seeing one's hand pointing at a painting during a museum tour or an animal in a zoo, relies on a multisensory integration of proprioceptive, visual, and tactile cues (Blanke et al., 2015). Subjective experiences create a link between episodic autobiographical memory and bodily self-consciousness, which suggests that the multisensory bodily signals may also be relevant to the conscious re-experiencing of self-relevant, past events.

Using virtual reality (VR) technology, Bergouignan et al. (2014) tested the encoding of real-life events from within one's own body/first-person perspective compared to outside one's own body/third-person perspective. Interestingly, the results showed episodic recollection deficits, specific to the events encoded in the outside body condition, associated with a diminished hippocampal activity (see **Table 1**). In a follow-up study (Bergouignan et al., 2021), the authors showed that encoding the real-life events from outside one's own body led to more third-person perspective during the retrieval (see **Table 1**). Recently, we tested whether the congruent multisensory bodily cues, i.e., the presence or absence of one's own physical body seen from a first-person perspective, would impact episodic autobiographical memory performance (**Figure 1A**). We used VR technology to create well-controlled, real life-like scenes into which participants, including their physical bodies, were immersed during the initial stage of encoding and later retrieval (Bréchet et al., 2019). We established that the presence of one's own physical body during encoding enhanced memory recognition and that this effect was body-specific (see **Table 1**). In a follow-up study by Gauthier et al. (2020), we demonstrated that seeing one's own body during encoding impacts the brain mechanisms responsible for episodic autobiographical memory formation by modulating the connectivity between the right hippocampal formation and the neocortical regions involved in the process of multisensory bodily signals and self-consciousness (see **Table 1**). In line with the previous work, Tacikowski et al. (2020) created an illusion of swapping participant's own body with a friend's body and then asked participants to perform personality rating and memory recognition tasks. The authors hypothesized that the perception of one's own body (bodily-self) would influence beliefs about one's own personality (conceptual-self) and that a coherent self-representation would lead to normal memory encoding. Indeed, they found that the experience of illusory ownership of a friend's body changed participants' beliefs about their own personality and made them more similar to the friend's personality. Interestingly, they further showed that adjusting to the new bodily-self was beneficial for memory encoding, while incoherence between the bodily-self and conceptual-self leads to memory impairment (see **Table 1**). The findings from these recent studies point out that memory retrieval is impaired when (i) participants encode events from outside their own bodies, (ii) the body is absent during encoding, and (iii) ownership of one's own body is reduced during encoding. These novel insights show that a coherent, multisensory representation of one's own body leads to hippocampal binding mechanisms that are directed to the neocortical areas, which are involved in episodic autobiographical memory (Bergouignan et al., 2014;

TABLE 1 | Studies examining the relationship between bodily-self and autobiographical-self.

Publications	How the aspects of the bodily-self influence the autobiographical self?
Bergouignan et al. (2014)	Out-of-body encoding causes episodic recollection deficits, associated with diminished hippocampal activity
Bergouignan et al. (2021)	Out-of-body encoding leads to more third-person perspective during recollection
Bréchet et al. (2018)	Brain activity related to self-location and 1PP anatomically overlap with episodic ABMs
Bréchet et al. (2019)	Seeing one;s own body during encoding enhances memory recognition
Bréchet et al. (2020)	Body-related integration is important for recall of episodic ABMs and prevents the loss of past events
Gauthier et al. (2020)	Seeing one's own body during encoding modulates connectivity between hippocampus and neocortical regions
Iriye and St Jacques (2020)	1 PP engages ABM retrieval network (i.e., hippocampus, anterior and posterior midline, frontal and posterior cortices) more strongly than 3 PP
Marcotti and St Jacques (2018)	Shifting visual perspective reduced the accuracy of subsequent memories
Penaud et al. (2022)	Familiarity and self-perspective improve recall and recognition of past events, their spatiotemporal context and sense of remembering
Piolino et al. (2009)	Re-experiencing past events through a feeling of self-awareness and 1 PP is are prone to fading over time
St Jacques et al. (2017)	Shifting visual perspective during ABM retrieval reshapes the characteristics of memories
St Jacques et al. (2018)	Remembering ABMs becomes more like imagination when shifting visual perspective
Tacikowski et al. (2020)	Self-concept can be updated by bodily-self changes; increase in self-coherence facilitates memory encoding

Gauthier et al., 2020; Iriye and St Jacques, 2020; Roehri et al., 2022).

BODILY-SELF AND AUTOBIOGRAPHICAL-SELF REPRESENTATIONS ANATOMICALLY OVERLAP

Until recently, it has been unknown whether bodily self-consciousness and autobiographical memories, either episodic or semantic, involve distinct or similar brain regions. We examined whether experimental manipulation of the self-location and first-person perspective aspects of bodily self-consciousness may anatomically overlap with the activations related to the subjective conscious experience of remembering self-relevant past events (Bréchet et al., 2018). We included results from patients suffering from out-of-body experiences with abnormal self-location and first-person perspective, whose brain damage was localized in the inferior parietal lobule (Ionta et al., 2011). Our systematic meta-analysis of neuroimaging studies revealed an anatomical overlap bilaterally in the angular gyrus, specific to bodily-self and episodic autobiographical-self, while there was no overlap with semantic autobiographical-self (see **Figure 1B** and **Table 1**). This finding is supported by the emerging evidence from studies in healthy participants that shows that the lateral parietal cortex is critical for the subjective, conscious experience of retrieving multisensory episodic memories (Bonnici et al., 2016, 2018; Sestieri et al., 2017; Humphreys et al., 2021). Related to this, recent series of papers (St Jacques et al., 2017, 2018; Marcotti and St Jacques, 2018; Iriye and St Jacques, 2020) showed that shifting visual perspective during retrieval shapes accuracy and subjective vividness of memories, which activated the posterior parietal cortex (see **Table 1**). Emerging evidence from patients with lateral parietal lesions shows that damage to this brain region causes reduced subjective confidence, vividness, and richness when re-experiencing self-relevant memories (Simons et al., 2010; Berryhill, 2012; Hower et al., 2014) as well as egocentric episodic memory deficits (Russell et al., 2019). Recently, Penaud et al. (2022) showed that familiarity and self-perspective (i.e., centered on one's own interaction with the environment) improved recall of past events, their spatiotemporal context and one's sense of remembering (see **Table 1**).

BODILY-SELF RETROACTIVELY STRENGTHENS AUTOBIOGRAPHICAL-SELF

Many seemingly irrelevant everyday life events may become significant only later (Kensinger, 2015). For example, that stranger who asked for directions becomes more relevant after realizing that your wallet is missing. Recently, two behavioral studies showed how memory for neutral images could be enhanced by future fearful (Dunsmoor et al., 2015) or rewarding (Patil et al., 2017), conceptually related events. More precisely, during the first stage of incidental encoding, the so-called *"pre-conditioning classification task,"* two neutral categories of images portraying animals and tools appeared to be of the same relevance. During a second phase of the incidental encoding, the so-called *"conditioning classification task,"* a prominent event, either fear conditioning or reward motivation, became purposefully associated with one of the two categories (animals or tools). A memory recognition task revealed that participants remembered better neutral images (for example, tools) associated with fear or reward during the conditioning phase. They also remembered better the conceptually related images (tools) from the pre-conditioning phase.

Studies on mental self-projections (Arzy et al., 2008; Dafni-Merom and Arzy, 2020) suggest that the experience of the self in the present moment is also involved with the ability to remember our past or imagine the future. Furthermore, this conscious self-awareness (i.e., a mental state in which the content of one's consciousness refers to knowledge about oneself, for example,

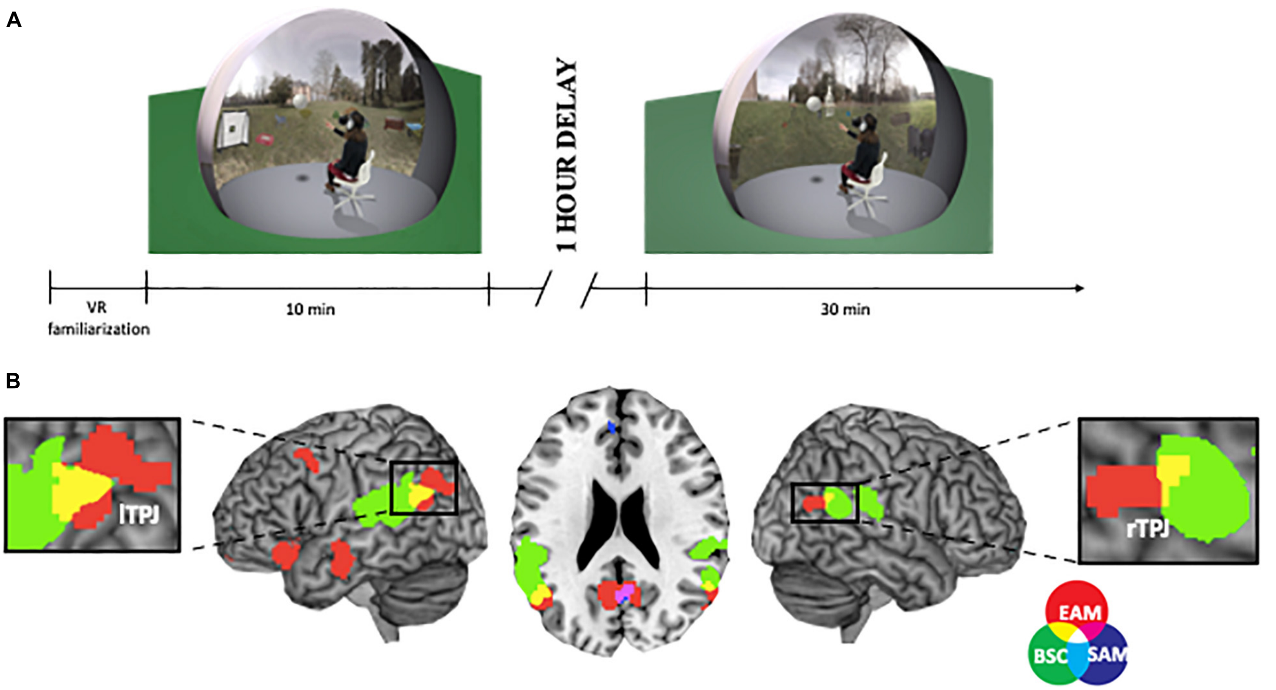


FIGURE 1 | (A) Study paradigm. First, participants incidentally learn the context of two different outside scenes (i.e., encoding session; 10 min). Participants were immersed back into the scenes with 1-h delay and were asked to perform a recognition task and subjective confidence ratings for each presented scene (i.e., retrieval session; 30 min). Before the actual study, subjects were seated in a chair, they were asked to put on the HMD and noise-canceling headphones to avoid external disturbances and familiarized themselves with the VR technology (10 min). **(B)** Visualization of the anatomical overlap between BSC fMRI analysis (i.e., self-location and first-person perspective) and EAM ALE analysis bilaterally in the parietal areas Label provided using the MRIcron. Within-cluster FEW-corrected $p < 0.05$ with $p < 0.001$ (uncorrected) as the cluster forming threshold.

reflecting about one's personality or identity) is intrinsically connected to the multisensory bodily processes (Blanke et al., 2015; Tacikowski et al., 2020). Therefore, we examined whether the retroactive and selective effect could be (a) triggered by the multisensory bodily signals, such as the presence or absence of one's own physical body and (b) generalized to naturalistic scenes, such as inside rooms or outdoor scenes, into which the participants would be immersed using VR technology (Bréchet et al., 2020). We showed that the presence of one's own body can retroactively strengthen memory recognition and that this retroactively enhancing effect became selectively associated with a particular group of items (either from rooms or scenes) (see **Table 1**).

SUMMARY AND FUTURE DIRECTIONS

The behavioral and neuroimaging evidence reviewed here suggests that the fundamental aspects of self-consciousness, the bodily-self and autobiographical-self, critically interact and impact each other. We experience the world from an inside viewpoint of a body and from a physical location of a body, which we identify as our own. This sense of ownership that I am the one who currently experiences the world around me is also essential with respect to one's past and future, therefore playing a significant role in constructing an autobiographical self. However, only a handful of recent studies (see **Table 1**) examined the relationship between bodily-self and autobiographical-self experimentally and showed that encoding events (i) from outside one's own body, (ii) when one's body is absent, (iii) when the sense of ownership of one's body is reduced leads to impairment of episodic autobiographical memories. On the other hand, these studies also revealed the coherent multisensory bodily-self representations have beneficial, strengthening effects on autobiographical memory and may prevent memory loss.

Understanding the interactions between the core aspects of self-consciousness, bodily-self, and autobiographical-self has important clinical implications. Several studies have found that memory decline in aging is associated with a lack of vivid autobiographical memories and increased retrieval from a third-person perspective (Piolino et al., 2006, 2009). Both normal aging and Alzheimer's disease in autobiographical memory are highly related to the self (Martinelli and Piolino, 2009; Martinelli et al., 2013). Alzheimer's disease is a neurodegenerative, progressive disorder that distorts the autobiographical-self, which is tightly connected to the sense of agency ("I am the one who generates experiences") and ownership ("I am the one who undergoes experiences") in the world (Arzy and Schacter, 2019). Detecting these early bodily-self signs, which are often overlooked in Alzheimer's disease, is of importance. One of the challenges for future research will be to reinstate a coherent sense of self to reduce autobiographical memory impairment.

Several open questions arise from these recent studies and their results. How can the neuroscientific methods ensure that people selectively remember what's significant to them in real life rather than laboratory-based settings? How can we ensure that only some memories are enhanced while others are diminished? Would those mentioned above, experimental studies tested in healthy young participants, modify the bodily-self and autobiographical-self of Alzheimer's patients in the same way? Is this the first step toward *"a memory pacemaker"*? More work is needed toward the long-term vision of restoring memory function. And perhaps more philosophically: What is more important: bodily-self or autobiographical-self? Quoting the words of Wiesel (2012): *"Illness may diminish me, but it will not destroy me. The body is not eternal, but the idea of the soul is. The brain will be buried, but memory will survive it."*

AUTHOR CONTRIBUTIONS

The author confirms being the sole contributor of this work and has approved it for publication.

FUNDING

This work was supported by Synapsis Foundation – Alzheimer Research Switzerland ARS to LB (Career Development Award).

ACKNOWLEDGMENTS

The author thank Daniel Schacter for his pertinent comments and suggestions on this manuscript.

REFERENCES

Arzy, S., Molnar-Szakacs, I., and Blanke, O. (2008). 'Self in time: imagined self-location influences neural activity related to mental time travel'. *J. Neurosci.* 28, 6502–6507. doi: 10.1523/JNEUROSCI.5712-07.2008

Arzy, S., and Schacter, D. L. (2019). 'Self-Agency and Self-Ownership in Cognitive Mapping'. *Trends Cogn. Sci.* 23, 476–487. doi: 10.1016/j.tics.2019.04.003

Bergouignan, L., Nyberg, L., and Ehrsson, H. H. (2014). 'Out-of-body-induced hippocampal amnesia'. *Proc. Natl. Acad. Sci. U.S.A.* 111, 4421–4426. doi: 10.1073/pnas.1318801111

Bergouignan, L., Nyberg, L., and Henrik Ehrsson, H. (2021). Out-of-body memory encoding causes third-person perspective at recall. *J. Cogn. Psychol.* 34, 160–178.

Berryhill, M. E. (2012). 'Insights from neuropsychology: pinpointing the role of the posterior parietal cortex in episodic and working memory'. *Front. Integr. Neurosci.* 6:31. doi: 10.3389/fnint.2012.00031

Blanke, O., Slater, M., and Serino, A. (2015). 'Behavioral, Neural, and Computational Principles of Bodily Self-Consciousness. *Neuron* 88, 145–166. doi: 10.1016/j.neuron.2015.09.029

Bonnici, H. M., Cheke, L. G., Green, D. A. E., FitzGerald, T., and Simons, J. S. (2018). Specifying a Causal Role for Angular Gyrus in Autobiographical Memory. *J. Neurosci.* 38, 10438–10443. doi: 10.1523/JNEUROSCI.1239-18.2018

Bonnici, H. M., Richter, F. R., Yazar, Y., and Simons, J. S. (2016). 'Multimodal Feature Integration in the Angular Gyrus during Episodic and Semantic Retrieval'. *J. Neurosci.* 36, 5462–5471. doi: 10.1523/JNEUROSCI.4310-15.2016

Bréchet, L., Grivaz, P., Gauthier, B., and Blanke, O. (2018). 'Common Recruitment of Angular Gyrus in Episodic Autobiographical Memory and Bodily Self-Consciousness'. *Front. Behav. Neurosci.* 12:270. doi: 10.3389/fnbeh.2018.00270

Bréchet, L., Hausmann, S. B., Mange, R., Herbelin, B., Blanke, O., and Serino, A. (2020). 'Subjective feeling of re-experiencing past events using immersive virtual reality prevents a loss of episodic memory'. *Brain Behav.* 10:e01571. doi: 10.1002/brb3.1571

Bréchet, L., Mange, R., Herbelin, B., Theillaud, Q., Gauthier, B., Serino, A., et al. (2019). 'First-person view of one's body in immersive virtual reality: influence on episodic memory'. *PLoS One* 14:e0197763. doi: 10.1371/journal.pone.0197763

Dafni-Merom, A., and Arzy, S. (2020). 'The radiation of autonoetic consciousness in cognitive neuroscience: a functional neuroanatomy perspective'. *Neuropsychologia* 143:107477. doi: 10.1016/j.neuropsychologia.2020.107477

Damasio, A. (1999). *The Feeling of What Happens: Body, Emotion and the Making of Consciousness*. San Diego: Harcourt Brace.

Damasio, A. (2003). Feelings of emotion and the self. *Ann. N. Y. Acad. Sci.* 1001, 253–261.

Dunsmoor, J. E., Murty, V. P., Davachi, L., and Phelps, E. A. (2015). 'Emotional learning selectively and retroactively strengthens memories for related events'. *Nature* 520, 345–348. doi: 10.1038/nature14106

Ehrsson, H. H. (2007). 'The experimental induction of out-of-body experiences'. *Science* 317:1048. doi: 10.1126/science.1142175

Gauthier, B., Bréchet, L., Lance, F., Mange, R., Herbelin, B., Faivre, N., et al. (2020). 'First-person body view modulates the neural substrates of episodic memory and autonoetic consciousness: a functional connectivity study'. *Neuroimage* 223:117370. doi: 10.1016/j.neuroimage.2020.117370

Hassabis, D., Kumaran, D., Vann, S. D., and Maguire, E. A. (2007). 'Patients with hippocampal amnesia cannot imagine new experiences'. *Proc. Natl. Acad. Sci. U.S.A.* 104, 1726–1731. doi: 10.1073/pnas.0610561104

Hower, K. H., Wixted, J., Berryhill, M. E., and Olson, I. R. (2014). 'Impaired perception of mnemonic oldness, but not mnemonic newness, after parietal lobe damage'. *Neuropsychologia* 56, 409–417. doi: 10.1016/j.neuropsychologia.2014.02.014

Humphreys, G. F., Lambon Ralph, M. A., and Simons, J. S. (2021). 'A Unifying Account of Angular Gyrus Contributions to Episodic and Semantic Cognition'. *Trends Neurosci.* 44, 452–463. doi: 10.1016/j.tins.2021.01.006

Ionta, S., Heydrich, L., Lenggenhager, B., Mouthon, M., Fornari, E., Chapuis, D., et al. (2011). 'Multisensory mechanisms in temporo-parietal cortex support self-location and first-person perspective'. *Neuron* 70, 363–374. doi: 10.1016/j.neuron.2011.03.009

Iriye, H., and St Jacques, P. L. (2020). 'How visual perspective influences the spatiotemporal dynamics of autobiographical memory retrieval'. *Cortex* 129, 464–475. doi: 10.1016/j.cortex.2020.05.007

Iriye, H., and St. Jacques, P. L. (2021). Memories for third-person experiences in immersive virtual reality. *Sci. Rep.* 11:4667. doi: 10.1038/s41598-021-84047-6

James, W. (1890). *The Principles of Psychology*. New York: Dover Publications.

Kensinger, E. A. (2015). 'The future can shape memory for the present'. *Trends Cogn. Sci.* 19, 179–180. doi: 10.1016/j.tics.2015.02.008

Lenggenhager, B., Tadi, T., Metzinger, T., Blanke, O., Ehrsson, H. H., Wiech, K., et al. (2007). 'Video ergo sum: manipulating bodily self-consciousness. Threatening a rubber hand that you feel is yours elicits a cortical anxiety response. *Science* 317, 1096–1099.

Mach, E. (1897). *The Analysis of Sensations and the Relation of the Physical to the Psychical*. La Salle, IL: Open Court.

Marcotti, P., and St Jacques, P. L. (2018). 'Shifting visual perspective during memory retrieval reduces the accuracy of subsequent memories'. *Memory* 26, 330–341. doi: 10.1080/09658211.2017.1329441

Martinelli, P., Anssens, A., Sperduti, M., and Piolino, P. (2013). 'The influence of normal aging and Alzheimer's disease in autobiographical memory highly related to the self'. *Neuropsychology* 27, 69–78. doi: 10.1037/a0030453

Martinelli, P., and Piolino, P. (2009). '[Self-defining memories: last episodic memories bastion in normal aging?]'. *Psychol. Neuropsychiatr. Vieil.* 7, 151–167. doi: 10.1684/pnv.2009.0178

Nielson, D. M., Smith, T. A., Sreekumar, V., Dennis, S., and Sederberg, P. B. (2015). 'Human hippocampus represents space and time during retrieval of real-world memories'. *Proc. Natl. Acad. Sci. U.S.A.* 112, 11078–11083. doi: 10.1073/pnas.1507104112

Nigro, G., and Neisser, U. (1983). 'Point of view in personal memories'. *Cogn. Psychol.* 15, 467–482. doi: 10.1080/09658211.2015.1137323

Park, H. D., and Blanke, O. (2019). 'Coupling Inner and Outer Body for Self-Consciousness'. *Trends Cogn. Sci.* 23, 377–388. doi: 10.1016/j.tics.2019.02.002

Patil, A., Murty, V. P., Dunsmoor, J. E., Phelps, E. A., and Davachi, L. (2017). 'Reward retroactively enhances memory consolidation for related items'. *Learn Mem.* 24, 65–69. doi: 10.1101/lm.042978.116

Penaud, S., Jebara, N., Zaoui, M., Orriols, E., Berthoz, A., and Piolino, P. (2022). 'Episodic memory and self-reference in a naturalistic context: new insights based on a virtual walk in the Latin Quarter of Paris'. *J. Environ. Psychol.* 81:101801.

Piolino, P., Desgranges, B., Clarys, D., Guillery-Girard, B., Taconnat, L., Isingrini, M., et al. (2006). 'Autobiographical memory, autonoetic consciousness, and self-perspective in aging'. *Psychol. Aging* 21, 510–525. doi: 10.1037/0882-7974.21.3.510

Piolino, P., Desgranges, B., and Eustache, F. (2009). 'Episodic autobiographical memories over the course of time: cognitive, neuropsychological and neuroimaging findings'. *Neuropsychologia* 47, 2314–2329. doi: 10.1016/j.neuropsychologia.2009.01.020

Roehri, N., Bréchet, L., Seeber, M., Pascual-Leone, A., and Michel, C. M. (2022). 'Phase-Amplitude Coupling and Phase Synchronization Between Medial Temporal, Frontal and Posterior Brain Regions Support Episodic Autobiographical Memory Recall'. *Brain Topogr.* 35, 191–206. doi: 10.1007/s10548-022-00890-4

Rubin, D. C., and Umanath, S. (2015). 'Event memory: a theory of memory for laboratory, autobiographical, and fictional events'. *Psychol. Rev.* 122, 1–23. doi: 10.1037/a0037907

Russell, C., Davies, S., Li, K., Musil, A. S., Malhotra, P. A., and Williams, A. L. (2019). 'Self-perspective in episodic memory after parietal damage and in healthy ageing'. *Neuropsychologia* 124, 171–181. doi: 10.1016/j.neuropsychologia.2018.12.013

Schacter, D. L., Addis, D. R., Hassabis, D., Martin, V. C., Spreng, R. N., and Szpunar, K. K. (2012). 'The future of memory: remembering, imagining, and the brain'. *Neuron* 76, 677–694. doi: 10.1016/j.neuron.2012.11.001

Sestieri, C., Shulman, G. L., and Corbetta, M. (2017). 'The contribution of the human posterior parietal cortex to episodic memory'. *Nat. Rev. Neurosci.* 18, 183–192. doi: 10.1038/nrn.2017.6

Simons, J. S., Peers, P. V., Mazuz, Y. S., Berryhill, M. E., and Olson, I. R. (2010). 'Dissociation between memory accuracy and memory confidence following bilateral parietal lesions'. *Cereb. Cortex* 20, 479–485. doi: 10.1093/cercor/bhp116

St Jacques, P. L., Carpenter, A. C., Szpunar, K. K., and Schacter, D. L. (2018). Remembering and imagining alternative versions of the personal past. *Neuropsychologia* 110, 170–179. doi: 10.1016/j.neuropsychologia.2017.06.015

St Jacques, P. L., and Schacter, D. L. (2013). Modifying memory: selectively enhancing and updating personal memories for a museum tour by reactivating them. *Psychol. Sci.* 24, 537–543. doi: 10.1177/0956797612457377

St Jacques, P. L., Szpunar, K. K., and Schacter, D. L. (2017). Shifting visual perspective during retrieval shapes autobiographical memories. *Neuroimage* 148, 103–114. doi: 10.1016/j.neuroimage.2016.12.028

Tacikowski, P., Weijs, M. L., and Ehrsson, H. H. (2020). 'Perception of Our Own Body Influences Self-Concept and Self-Incoherence Impairs Episodic Memory. *iScience* 23:101429. doi: 10.1016/j.isci.2020.101429

Tulving, E. (1985). 'Memory and consciousness'. *Can. Psychol.* 26, 1–12.

Vogel, S., and Schwabe, L. (2016). 'Stress in the zoo: tracking the impact of stress on memory formation over time'. *Psychoneuroendocrinology* 71, 64–72. doi: 10.1016/j.psyneuen.2016.04.027

Wiesel, E. (2012). *Open Heart*. New York: Alfred A. Knopf.

Zaman, A., and Russell, C. (2021). Does autonoetic consciousness in episodic memory rely on recall from a first-person perspective? *J. Cogn. Psychol*. 34, 9–23.

Conflict of Interest: The author declares that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's Note: All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

frontiers | Frontiers in Psychology

TYPE Hypothesis and Theory
PUBLISHED 21 October 2022
DOI 10.3389/fpsyg.2022.1008407

Check for updates

OPEN ACCESS

EDITED BY
Antonino Raffone,
Sapienza University of Rome, Italy

REVIEWED BY
Pablo Villegas,
Enrico Fermi Center for Study and Research, Italy
Fabrizio Lombardi,
ETH Zürich, Switzerland

*CORRESPONDENCE
Giovanni Rabuffo
giovanni.rabuffo@univ-amu.fr

SPECIALTY SECTION
This article was submitted to Consciousness Research, a section of the journal Frontiers in Psychology

RECEIVED 31 July 2022
ACCEPTED 04 October 2022
PUBLISHED 21 October 2022

CITATION
Rabuffo G, Sorrentino P, Bernard C and Jirsa V (2022) Spontaneous neuronal avalanches as a correlate of access consciousness.
Front. Psychol. 13:1008407.
doi: 10.3389/fpsyg.2022.1008407

Spontaneous neuronal avalanches as a correlate of access consciousness

Giovanni Rabuffo*, Pierpaolo Sorrentino, Christophe Bernard and Viktor Jirsa

Institut de Neurosciences des Systemes, Aix-Marseille University, Marseille, France

Decades of research have advanced our understanding of the biophysical mechanisms underlying consciousness. However, an overarching framework bridging between models of consciousness and the large-scale organization of spontaneous brain activity is still missing. Based on the observation that spontaneous brain activity dynamically switches between epochs of segregation and large-scale integration of information, we hypothesize a brain-state dependence of conscious access, whereby the presence of either segregated or integrated states marks distinct modes of information processing. We first review influential works on the neuronal correlates of consciousness, spontaneous resting-state brain activity and dynamical system theory. Then, we propose a test experiment to validate our hypothesis that conscious access occurs in aperiodic cycles, alternating windows where new incoming information is collected but not experienced, to punctuated short-lived integration events, where conscious access to previously collected content occurs. In particular, we suggest that the integration events correspond to neuronal avalanches, which are collective bursts of neuronal activity ubiquitously observed in electrophysiological recordings. If confirmed, the proposed framework would link the physics of spontaneous cortical dynamics, to the concept of ignition within the global neuronal workspace theory, whereby conscious access manifest itself as a burst of neuronal activity.

KEYWORDS

consciousness, neuronal avalanches, criticality, conscious access, resting state, spontaneous activity

1. Introduction

The origin of subjective lived experience, generally referred to as Consciousness, poses some of the most fascinating scientific and philosophical questions (Nagel, 1974). The advances of paradigms from the field of cognitive psychology, as well as theoretical models from computational neuroscience, have been pushing for refined definitions of Consciousness, and for a unified framework to understand and guide the interpretation of accumulating experimental evidences (Wiese, 2020; Melloni et al., 2021; Signorelli et al., 2021). A major distinction between "phenomenal" and "access" Consciousness was proposed (Block, 2005). Phenomenal (P-)Consciousness refers to raw experience (qualia; Chalmers, 1995). Access (A-)Consciousness refers to the availability of information for explicit reasoning and rational control. The latter definition allowed precise dissection of brain activity upon sensory stimuli that are accompanied by consciousness vs. stimuli that are not. In particular, task-based neuroimaging experiments led to the identification of robust neuronal correlates of A-Consciousness (NCC; Dehaene and Changeux, 2011; Aru et al., 2012). Generally speaking, event-related potentials (ERPs) display a subliminal early response in sensory areas regardless of the presence of Consciousness, then followed by characteristic waves of activity whose magnitude marks the presence or absence of perceptual awareness. For example, visual awareness negativity observed 200 ms after a visual stimulus, and enhanced P3 amplitude observed after 300 ms are often considered candidate NCCs (Koivisto and Revonsuo, 2010; Salti et al., 2012; Koivisto and Grassini, 2016). The Global Neuronal Workspace (GNW) hypothesis describes the crossing of the subliminal threshold and the consequent wave of activation as a "global ignition" (Mashour et al., 2020), which is postulated to be necessary for conscious access. These experiments elegantly identified the correlates of the conscious experience elicited by artificial stimuli. However, the required controlled experimental settings are far from naturalistic scenarios, where external and endogenous stimuli occur across multiple spatial and temporal scales. Another way of studying consciousness is by investigating the organizational principles of brain activity in altered states of consciousness. Studying functional imaging in such conditions led to the observation that the capability of supporting consciousness goes along with increasingly complex brain activities, which can be evaluated by using signal diversity measures such as Lempel-Ziv complexity (Casali et al., 2013; Schartner et al., 2015; Arena et al., 2021). In fact, the degree of complexity in brain signals provides a new mean to assess the presence of consciousness in a clinical setting, regardless of patient responsiveness (Sanders et al., 2012). The use of complexity measures has been influenced by the Integrated Information Theory (IIT; Oizumi et al., 2014; Tononi et al., 2016), which proposed the scalar quantity Φ as a measure of the quantity and quality of conscious experience, which concerns P-consciousness. In fact, it is possible to interpret Φ as a measure of information-processing complexity as well as dynamical systems complexity (Mediano et al., 2022). From a neurodynamical point of view, consciousness was associated to the brain being able to spontaneously explore a rich dynamical repertoire of network states, whereas unconscious states were associated to a less complex network dynamics (Uhrig et al., 2018; Demertzi et al., 2019). In fact, even at rest, when the brain is not involved in any specific task, large-scale brain activity dynamically organizes in communities of strongly correlated brain regions, or resting state networks, as observed by dynamic Functional Connectivity measures (Preti et al., 2017). Notably, resting state networks do not evolve continuously, nor periodically, but rather in aperiodic bursts of network co-fluctuations (Tagliazucchi et al., 2012; Zamani Esfahlani et al., 2020; Rabuffo et al., 2021). Similarly, large scale brain activity is often described as alternating segregated moments, in which regional activities are prominently independent from each other, to integrated moments when large-scale interaction occurs (Sporns, 2013; Deco et al., 2015; Shine, 2019).

While the importance of such dynamic reconfiguration for brain function is generally recognized (Lord et al., 2017), it is not clear how the spontaneous switching between segregated and integrated states relates to A- and P- Consciousness. For example, let us suppose that a visual stimulus is flashed to our retina. Does the presence of either brain state (i.e., segregated or integrated) affect the probability of such stimulus to gain conscious access? In this work we propose that information is collected predominantly during the segregated state, and that part of such information gains conscious access at the subsequent integrated state (see Section 4). In particular, we identify a fine-grained correlate of A-Consciousness corresponding to large-scale bursts of neuronal activations, interpreted as integration events. In the next chapter we associate these salient events to neuronal avalanches, as understood in the context of "brain criticality" theory. We review recent developments in Consciousness studies within this framework, we draw postulates of our hypotheses and propose experiments to test them.

2. Avalanches and consciousness

Consciousness should be understood within the physical principles governing the brain (Cosmelli et al., 2007; Werner, 2007). A widely discussed hypothesis is that, likewise other complex systems outside of equilibrium, the brain self regulates around a critical point i.e., at the edge of a second-order phase transition (Cocchi et al., 2017; OByrne and Jerbi, 2022), a property that is known as Self-Organized Criticality (SOC; Bak et al., 1987; Plenz et al., 2021). SOC offers an attractive theoretical framework for the brain, since it predicts the empirical evidences of optimal information processing (Shew et al., 2011; Shew and Plenz, 2013), dynamical range (Kinouchi and Copelli, 2006; Larremore et al., 2011), and maximization of metastability (Tognoli and Kelso, 2014). It was previously suggested that SOC could serve as a framework for Consciousness as well (Werner, 2009; Carhart-Harris et al., 2014; Tagliazucchi, 2017; Walter and Hinterberger, 2022). In fact, criticality was linked to IIT as a necessary condition for integration of information (Kim and Lee, 2019; Popiel et al., 2020), and it is compatible with predictions from the GNW hypothesis (Tagliazucchi, 2017), among other frameworks. The typical signature of SOC is the presence of neuronal avalanches, corresponding to sudden chains of neuronal activations across the brain. Neuronal avalanches are typically characterized by their duration—up to a few hundreds milliseconds—and their size i.e., the number of regions recruited. The distribution of the sizes of the avalanches follows a power-law, which indicates that these events span several orders of magnitude. In fact, neuronal avalanches can be consistently observed at the neuronal level using local multielectrode arrays (Beggs and Plenz, 2003), as well as at the whole-brain level using EEG (Palva et al., 2013), MEG (Shriki et al., 2013), SEEG (Priesemann et al., 2013), and fMRI (Tagliazucchi et al., 2012). It is to be noted that the outburst of an avalanche corresponds to the strong co-fluctuation of a set of brain regions, which promotes resting-state network dynamics (Tagliazucchi et al., 2012; Zamani Esfahlani et al., 2020; Rabuffo et al., 2021), among other signal properties. We propose that large neuronal avalanches support large-scale integration and are a candidate NCC at a fine-grained temporal scale. Accordingly, our hypotheses also apply in conditions such as resting-state. In our framework conscious access to external stimuli, would depend on the relative timing of the stimulus with respect to the spontaneous background avalanche dynamics.

3. Localize consciousness

An open question is the location of the physical substrate of a conscious experience. While it is generally accepted that Consciousness involves a distributed process across the brain regions, a major dichotomy lies in the anterior or posterior localization (Boly et al., 2017; Odegaard et al., 2017). For example, it was suggested that content-specific NCCs lay in neuronal ensembles within a hot zone situated in posterior cortical regions (Koch et al., 2016). Other evidences suggest that the prefrontal cortex is fundamental for the ignition upon (and therefore the conscious access to) visual inputs (Joglekar et al., 2018; Van Vugt et al., 2018). How these studies explain consciousness in absence of experimentally controlled stimuli is still open to debate (Mashour, 2018). Furthermore, evoked activation patterns underlying a conscious percept depend on the nature of the stimulus e.g., visual vs. auditory (Eriksson et al., 2007). Our proposal of linking A-Consciousness to spontaneous neuronal avalanches allows to interpret these results in the framework of resting-state activity. Functional imaging at rest revealed that healthy brains are characterized by a high-number of network reconfigurations (brain flexibility), a property which is disrupted in neurodegenerative diseases (Sorrentino et al., 2021a). In the context of dynamical systems theory, brain flexibility derives from the multi-stability of the dynamical repertoire (Golos et al., 2015), which has been linked to cognition (Kelso, 2012). Being dynamical in nature, avalanches recruit several modules in a flexible way, acting as a physical manifestation of multi-stability. We propose that fine-grained NCCs are coded into ever-changing patterns of network activations, which manifest as neuronal avalanches. However, we argue that a certain threshold size and/or topographic boundary should exist for an avalanche in order to be relevant for conscious access (see Section 5). Hence, the question of the localization of consciousness is reinterpreted from a network perspective, whereby the role that individual regions play into the spread and reconfiguration of neuronal avalanches is not homogeneous. In fact, the probability of cascading along a specific brain network resembles the structural connectome (Sorrentino et al., 2021b), suggesting that avalanches follow preferential pathways laid out by the neuronal projections. Thus, the topological role of a region within the connectome will partly define its relevance toward the spreading of perturbations on the large-scale. This is in line with the fact that the effect of a lesion on the large-scale dynamics correlate very poorly with the size of the lesion, while the topological role of the lesion carries much higher predictive power (Gratton et al., 2012).

4. Brain-state dependence of consciousness

A recent study on monkeys suggested that conscious access is state-dependent, whereby fluctuations in the behavioral and neurophysiological markers before the visual stimulus were related to variations in stimulus detection (Van Vugt et al., 2018). Similarly, we expect a brain-state dependence of conscious access, given that the presence or absence of an avalanche mark qualitatively distinct states in the functional evolution of brain dynamics. In fact, neuronal ensembles are not always in communication, and windows of coherence are expected to be necessary for information routing (Fries, 2005). In particular, large-scale coordination displays the dynamic alternation of regional segregation and integration (Friston, 2009; Sporns, 2013; Deco et al., 2015). The sequential recruitment of neuronal ensembles during avalanches elicits characteristic brain networks, and corresponds to periods of increased functional connectivity. Thus, under the communication through coherence hypothesis (Fries, 2005), we assume that neuronal avalanches underlie periods of integration. As a first conceptualization of our framework, based on the idea that integration of information is a prerequisite for consciousness (Tononi et al., 1998a), we posit that avalanches spreading across the brain correspond to discrete conscious access events, where previously acquired information is broadcast in the brain. On the one hand, we argue that avalanche states are poor receptive states, since the already-entrained populations are generally less susceptible to external stimuli. On the other hand, if a stimulus is received during an inter-avalanche interval, it has a higher probability of triggering a new avalanche and of gaining conscious access after a few hundreds milliseconds. We can express these concepts in the language of dynamical systems and manifolds. In this representation, the state of the system can be thought of as a ball rolling over a surface (manifold), whose wells (attractors) represent quasi-stable configurations of the system (Figures 1A,B). The brain activity spends a large amount of time in one attractor, and then quickly transitions to

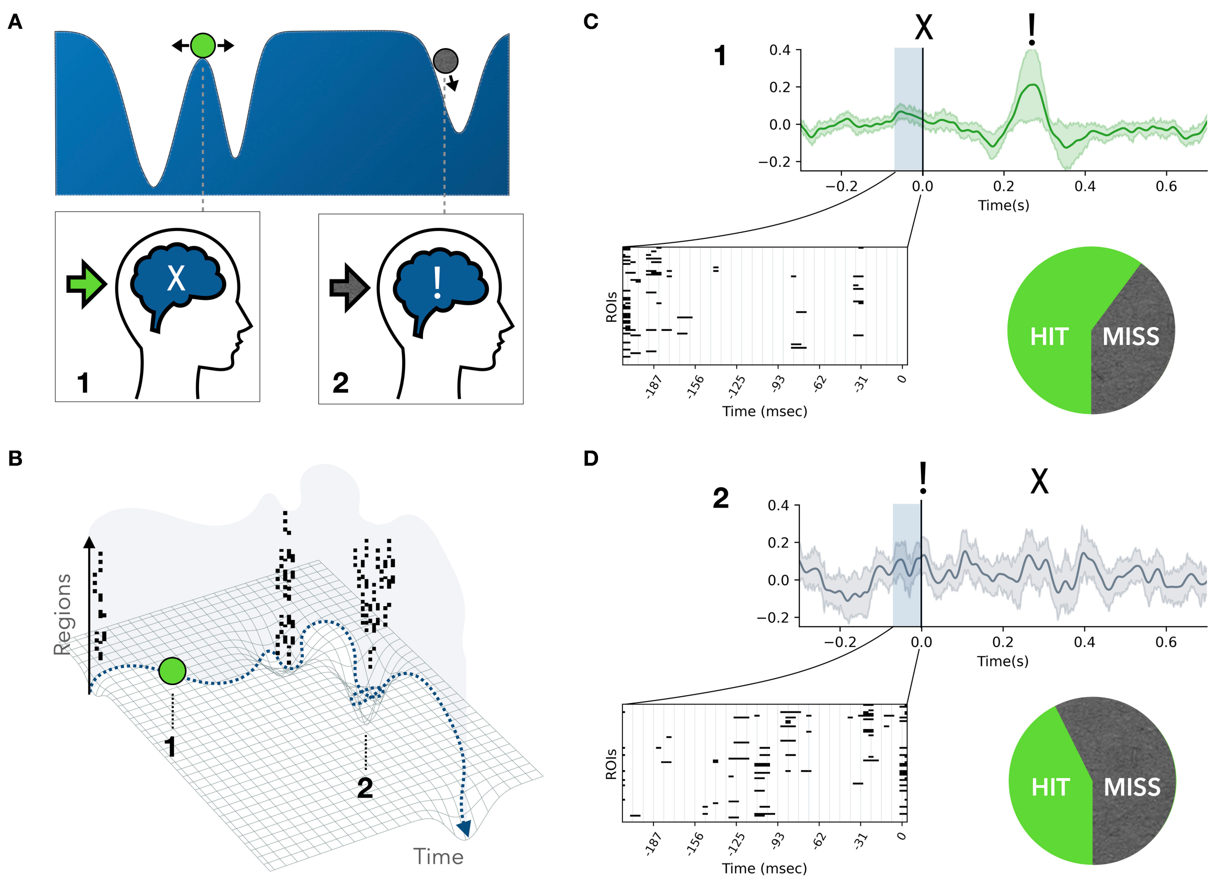


FIGURE 1
Conceptual framework. **(A,B)** The brain dynamics can be represented as a ball rolling on a landscape. During an inter-avalanche interval (state 1) the system explores the high plateau and can be easily pushed into a well by a perturbation. Therefore, in state 1 the system is maximally receptive to an incoming stimulus (green arrow), but not aware of it yet (marked by "X"). On the contrary, during an avalanche (state 2) the system is falling into a well, and an external perturbation has less impact on the system (gray arrow). An avalanche is characterized by the simultaneous fluctuation of several regions above a fixed threshold (marked by black bins). Therefore, state 2 supports the formation of a large-scale network of highly co-activated brain regions. We interpret this as a moment of integration of information across regions, where conscious access to previously collected information is achieved (marked as "!"). **(C,D)** Access to conscious processing occurs few hundred milliseconds after a stimulus is sent (at time "zero"), and this event is generally marked by a late wave of activations (ignition). Let us suppose that in an experimental setup a stimulus intensity is tuned such that perceptual awareness and the associated ignition occur in fifty percent of trials. We predict that if the stimulus is sent during an inter-avalanche interval **(C)**, the probability of gaining conscious access a few hundred milliseconds later is greater than fifty percent. Accordingly, the probability drops below fifty percent if the stimulus coincides with a large avalanche **(D)**.

the next one. Over time, multiple attractors (i.e., "the landscape") will have been visited. Hence, rather than stationary, the patterns of activity in the data will be *metastable* (Haken, 1991; Friston, 1997; Huys et al., 2014; Deco and Kringelbach, 2016; Roberts et al., 2019; Shine et al., 2019).

Neuronal avalanches correspond on the manifold to moments in which the ball is (falling) into a well (and, as said, conscious access to previously collected stimuli is gained). In this situation it is harder for a new incoming external perturbation (stimulus) to change the trajectory of the system. On the contrary, when the ball is on the peak, the system is more susceptible to external stimuli, which can easily push the ball into a new attractor (i.e., update the conscious percept). Summarizing, we suggest that the brain alternates receptive windows (inter-avalanche intervals), where incoming information is collected, to punctuated short-lived integration events (large avalanches), corresponding to conscious access to the previously collected information.

Finally, we propose a general experimental setting to test these hypotheses. In task-based experiments it is possible to fine-tune the duration and intensity of a sensory stimulus (e.g., visual) to a target subject-specific probability of such stimulus gaining conscious access. Our hypotheses predict an increased probability of missing the target if that is presented during a large avalanche. This can be tested by simultaneously recording the spontaneous avalanche state at the moment of the presentation of the stimulus (Figures 1C,D). However, given the aperiodic occurrence of avalanches, it is challenging to time the presentation of a stimulus to a specific brain state. To overcome the problem, we propose to exploit an important property of neuronal avalanches, namely the shape collapse (Sethna et al., 2001). When an avalanche breaks out, regions are gradually recruited until a maximum is reached, and then the activity fades away equally gradually. This phenomenon holds across orders of magnitude, which allows all avalanche profiles to be collapsed into a parabolic shape by a fixed scaling exponent, a property expected theoretically for a class of universality of critical phenomena (Papanikolaou et al., 2011; Laurson et al., 2013; Miller et al., 2019). This implies a short-term form of predictability over the course of an avalanche. In fact, once an avalanche reaches the maximum recruitment after N steps, one can predict that it will persist for the next N time steps, symmetrically collapsing in time, up to few hundred milliseconds for the longest avalanches. This would provide a time-window to synchronize the stimulus to the background avalanche state. Importantly, such probability can be evaluated in the absence of any task or behavior using a no-report paradigm, such as Sergent et al. (2021).

5. Discussion

The brain activity is functionally dynamic and conscious processes might depend on the spontaneous alternation of integrated and segregated functional states. Neuronal avalanches are ubiquitously observed in brain imaging recordings, in association with expected features underpinning Consciousness, such as complexity, flexibility and multistability. In this work, we propose an overarching framework to link a number of empirical findings related to consciousness and its neuronal correlates. Our hypotheses remain highly hypothetical. However, we propose an experimental design to test our framework. In detail, we propose that a computational cycle for consciousness consists of a subliminal phase, where the brain operates in a segregated state and external stimuli are gathered, and an integrated state manifesting as a neuronal avalanche, where the previously acquired information can gain conscious access. In our framework, conscious access manifest itself as large-scale bursts of activation. Such bursts can be induced by external stimuli (e.g., ignition), but are also empirically present in spontaneous resting-state activity (e.g., neuronal avalanches). Hence, our hypotheses ought to explain Consciousness also in absence of a clear task e.g., when we are mind-wandering. In general, in order to identify the necessary and sufficient conditions for an avalanche to support Consciousness, it is important to distinguish between neuronal avalanches generated during conscious or unconscious states. In this context, the topography of the networks recruited by an avalanche is expected to be relevant. In fact, specific neuronal circuits supporting the emergence and fading of Consciousness have been identified in the literature. As an example, the cortico-cortical and thalamo-cortical loops are associated to both states and contents of Consciousness (Dynamic Core hypothesis; Edelman and Tononi, 2000). At the mesoscale, the pyramidal neurons in layer 5 (L5p) in the cortex have been identified as a key relay between the cortico-cortical and the thalamo-cortical loops. As such, the involvement of L5p has been hypothesized as a necessary condition for cortical processes to support consciousness (Aru et al., 2019). Recent evidence highlight changes in L5p activity during anesthesia, such as increased low-frequency power (Bastos et al., 2021), and selective synchronization (Bharioke et al., 2022). Supposedly, these changes impair the dendritic-to-soma coupling (Aru et al., 2020; Suzuki and Larkum, 2020), thereby disconnecting the thalamo-cortical broadcasting system which would no longer support consciousness. In the light of this evidence, we hypothesize that a conscious state should be supported by avalanches that contribute to the integration of these systems, which is more likely for large-sized avalanches. However, we do not exclude that access consciousness can be supported by small avalanches, provided that they recruit the relevant structures. Classical statistical measures related to neuronal avalanches (e.g., power law distribution, branching ratio etc..), often used as a signature of (or against) criticality, might not be optimal to assess conscious vs. unconscious states as they disregard both the topography and the temporal organization of brain activity. The temporal factor is also crucial to the emerge of consciousness, as a large body of research suggests that Consciousness involves multiple characteristic timescales, spanning from the perceptive threshold ($15-50ms$) to the extension ($100-500ms$) and the retention ($3-7s$) of content (see Singhal and Srinivasan, 2021 and references within). While a number of frameworks describing the temporal hierarchy of conscious perception have been proposed (e.g., Pöppel, 1997; Varela, 1999; Poppel, 2004; Wanja, 2017), a clear biophysical mechanism encompassing fast, intermediate and slow temporal scales is still missing. To this regard, neuronal avalanches offer a promising overarching framework to explain the spontaneous emergence of this hierarchy of timescales. In fact, the duration of neuronal avalanches is rather short on average (tens of milliseconds; comparable to the perceptive threshold) with respect to the duration of IAI (hundreds of milliseconds; which is closer to the extensional timescales). Furthermore, it was shown that the dynamics of avalanches possesses a slower timescale in the order of a few seconds (Rabuffo et al., 2021), which is comparable to the retentional timescale. However, it must be noted that a separation of timescales predicted by the SOC framework is not clearly observed in experimental recordings (Lombardi et al., 2012, 2021; Priesemann et al., 2014). It was previously shown that large avalanches tend to couple with longer preceding and following IAI (Lombardi et al., 2016), which might be related to the proposed hypothesis of information gathering during IAI and conscious access during the following large avalanches (and also be of use to design the experiments). It had been proposed that cognition and perception might operate in discrete cycles (VanRullen and Koch, 2003; Madl et al., 2011; VanRullen, 2016). The identification of these cycles with recurrent neuronal avalanches suggests that fast cognitive processes might be aperiodic, and that the duration and intensity of a percept could be modulated in time (Herzog et al., 2016), likewise the duration and size of neuronal avalanches.

Arguably, one of the most used ontologies for the description of neuronal activity in neuroscience refers to neuronal oscillations (Buzsaki, 2006). In this framework, transient neuronal synchronization events are considered as a correlate of high-order cognitive processing (Tononi et al., 1998a,b; Rodriguez et al., 1999; Srinivasan et al., 1999; Engel and Singer, 2001; Ward, 2003; Melloni et al., 2007). In particular, synchronization of neuronal populations has been proposed to mediate the merging of multiple local processes into a single conscious experience (Singer, 2001). Hence, a mechanistic framework for consciousness should ideally portray both oscillations and avalanches. Recent works has provided both theoretical and experimental evidence in this direction, showing nested oscillations coexisting with neuronal avalanches (Gireesh and Plenz, 2008). In the same vein, it was shown *in-silico* that neuronal avalanches occur at a critical (asynchronous-to-synchronous) phase transition, where they coexist with incipient oscillations (Di Santo et al., 2018). However, while SOC provides a candidate framework for a unified theory of Consciousness (Melloni et al., 2021), a critical state is not strictly required for explaining the presence of spontaneous scale-free neuronal avalanches (e.g., Buendía et al., 2020a, 2022). For example, it was suggested that during deep sleep or under anesthesia, the brain self-organizes at the edge of bistability (SOB, a first order phase transition Buendía et al., 2020b), rather than SOC (Priesemann et al., 2013). Furthermore, other bursting phenomena such as EEG micro states (Britz et al., 2009; Michel and Koenig, 2018) and neuronal assemblies (Papadimitriou et al., 2020) have been linked to cognition and perception. Notably, the global ignition proposed by the GNW hypothesis and observed when an external stimulus gains conscious access, might be understood as an avalanche of neuronal activations.

In conclusion, in this manuscript we propose a unifying framework to link both spontaneous and induced conscious access to neuronal avalanches. Neuronal avalanches are a solid finding in large-scale human recordings, which allow to state testable hypotheses and design the experiments accordingly. The presence of ignition when a stimulus gains conscious access, as predicted by the GNW hypothesis, and the loss of complexity in unconscious states, predicted by IIT, can both be understood within the proposed framework. Importantly, neuronal avalanches also offer a mathematically solid approach to link the properties observed on the large-scale to hypothetical microscopic mechanisms, which might offer a window to approach consciousness in mechanistic terms.

Data availability statement

The original contributions presented in the study are included in the article/supplementary material, further inquiries can be directed to the corresponding author.

Author contributions

GR and PS conceived the presented ideas and contributed to the writing of the manuscript. GR designed the figure. CB and VJ supervised the findings and the writing of this work. All authors discussed and contributed to the final manuscript.

Funding

This work was supported by the grant ANR-17-CE37-0001-CONNECTOME, by the European Union's Horizon 2020 research and innovation programme under grant agreement No. 945539 (SGA3) Human Brain Project and VirtualBrainCloud No. 826421.

Conflict of interest

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

Publisher's note

References

Arena, A., Comolatti, R., Thon, S., Casali, A. G., and Storm, J. F. (2021). General anesthesia disrupts complex cortical dynamics in response to intracranial electrical stimulation in rats. *eNeuro* 8, ENEURO.0343-20.2021. doi: 10.1523/ENEURO.0343-20.2021

Aru, J., Bachmann, T., Singer, W., and Melloni, L. (2012). Distilling the neural correlates of consciousness. *Neurosci. Biobehav. Rev*. 36, 737–746. doi: 10.1016/j.neubiorev.2011.12.003

Aru, J., Suzuki, M., and Larkum, M. E. (2020). Cellular mechanisms of conscious processing. *Trends Cogn. Sci*. 24, 814–825. doi: 10.1016/j.tics.2020.07.006

Aru, J., Suzuki, M., Rutiku, R., Larkum, M. E., and Bachmann, T. (2019). Coupling the state and contents of consciousness. *Front. Syst. Neurosci*. 13, 43. doi: 10.3389/fnsys.2019.00043

Bak, P., Tang, C., and Wiesenfeld, K. (1987). Self-organized criticality: an explanation of the 1/f noise. *Phys. Rev. Lett*. 59, 381. doi: 10.1103/PhysRevLett.59.381

Bastos, A. M., Donoghue, J. A., Brincat, S. L., Mahnke, M., Yanar, J., Correa, J., et al. (2021). Neural effects of propofol-induced unconsciousness and its reversal using thalamic stimulation. *Elife* 10. doi: 10.7554/eLife.60824.sa2

Beggs, J. M., and Plenz, D. (2003). Neuronal avalanches in neocortical circuits. *J. Neurosci*. 23, 11167–11177. doi: 10.1523/JNEUROSCI.23-35-11167.2003

Bharioke, A., Munz, M., Brignall, A., Kosche, G., Eizinger, M. F., Ledergerber, N., et al. (2022). General anesthesia globally synchronizes activity selectively in layer 5 cortical pyramidal neurons. *Neuron* 110, 2024.e10–2040.e10. doi: 10.1016/j.neuron.2022.03.032

Block, N. (2005). Two neural correlates of consciousness. *Trends Cogn. Sci*. 9, 46–52. doi: 10.1016/j.tics.2004.12.006

Boly, M., Massimini, M., Tsuchiya, N., Postle, B. R., Koch, C., and Tononi, G. (2017). Are the neural correlates of consciousness in the front or in the back of the cerebral cortex? clinical and neuroimaging evidence. *J. Neurosci*. 37, 9603–9613. doi: 10.1523/JNEUROSCI.3218-16.2017

Britz, J., Landis, T., and Michel, C. M. (2009). Right parietal brain activity precedes perceptual alternation of bistable stimuli. *Cereb. Cortex* 19, 55–65. doi: 10.1093/cercor/bhn056

Buendía, V., Di Santo, S., Bonachela, J. A., and Mu noz, M. A. (2020a). Feedback mechanisms for self-organization to the edge of a phase transition. *Front. Phys*. 8, 333. doi: 10.3389/fphy.2020.00333

Buendía, V., di Santo, S., Villegas, P., Burioni, R., and Muñoz, M. A. (2020b). Self-organized bistability and its possible relevance for brain dynamics. *Phys. Rev. Res*. 2, 013318. doi: 10.1103/PhysRevResearch.2.013318

Buendía, V., Villegas, P., Burioni, R., and Mu noz, M. A. (2022). The broad edge of synchronization: griffiths effects and collective phenomena in brain networks. *Philos. Trans. R. Soc. A* 380, 20200424. doi: 10.1098/rsta.2020.0424

Buzsaki, G. (2006). *Rhythms of the Brain*. Oxford: Oxford University Press.

Carhart-Harris, R. L., Leech, R., Hellyer, P. J., Shanahan, M., Feilding, A., Tagliazucchi, E., et al. (2014). The entropic brain: a theory of conscious states informed by neuroimaging research with psychedelic drugs. *Front. Human Neurosci*. 8, 20. doi: 10.3389/fnhum.2014.00020

Casali, A. G., Gosseries, O., Rosanova, M., Boly, M., Sarasso, S., Casali, K. R., et al. (2013). A theoretically based index of consciousness independent of sensory processing and behavior. *Sci. Transl. Med*. 5, 198ra105. doi: 10.1126/scitranslmed.3006294

Chalmers, D. J. (1995). Facing up to the problem of consciousness. *J. Conscious. Stud*. 2, 200–219.

Cocchi, L., Gollo, L. L., Zalesky, A., and Breakspear, M. (2017). Criticality in the brain: a synthesis of neurobiology, models and cognition. *Progr. Neurobiol*. 158, 132–152. doi: 10.1016/j.pneurobio.2017.07.002

Cosmelli, D., Lachaux, J.-P., and Thompson, E. (2007). "Neurodynamical approaches to consciousness," in *The Cambridge Handbook of Consciousness*, eds P. D. Zelazo, M. Moscovitch, and E. Thompson (Cambridge University Press). doi: 10.1017/CBO9780511816789.027

Deco, G., and Kringelbach, M. L. (2016). Metastability and coherence: extending the communication through coherence hypothesis using a whole-brain computational perspective. *Trends Neurosci*. 39, 125–135. doi: 10.1016/j.tins.2016.01.001

Deco, G., Tononi, G., Boly, M., and Kringelbach, M. L. (2015). Rethinking segregation and integration: contributions of whole-brain modelling. *Nat. Rev. Neurosci*. 16, 430–439. doi: 10.1038/nrn3963

Dehaene, S., and Changeux, J.-P. (2011). Experimental and theoretical approaches to conscious processing. *Neuron* 70, 200–227. doi: 10.1016/j.neuron.2011.03.018

Demertzi, A., Tagliazucchi, E., Dehaene, S., Deco, G., Barttfeld, P., Raimondo, F., et al. (2019). Human consciousness is supported by dynamic complex patterns of brain signal coordination. *Sci. Adv*. 5, eaat7603. doi: 10.1126/sciadv.aat7603

Di Santo, S., Villegas, P., Burioni, R., and Mu noz, M. A. (2018). Landau-ginzburg theory of cortex dynamics: scale-free avalanches emerge at the edge of synchronization. *Proc. Natl. Acad. Sci. U.S.A*. 115, E1356-E1365. doi: 10.1073/pnas.1712989115

Edelman, G. M., and Tononi, G. (2000). "Reentry and the dynamic core: Neural correlates of conscious experience," in *Neural Correlates of Consciousness*, ed T. Metzinger (MIT Press).

Engel, A. K., and Singer, W. (2001). Temporal binding and the neural correlates of sensory awareness. *Trends Cogn. Sci*. 5, 16–25. doi: 10.1016/S1364-6613(00)01568-0

Eriksson, J., Larsson, A., Åhlström, K. R., and Nyberg, L. (2007). Similar frontal and distinct posterior cortical regions mediate visual and auditory perceptual awareness. *Cereb. Cortex* 17, 760–765. doi: 10.1093/cercor/bhk029

Fries, P. (2005). A mechanism for cognitive dynamics: neuronal communication through neuronal coherence. *Trends Cogn. Sci*. 9, 474–480. doi: 10.1016/j.tics.2005.08.011

Friston, K. J. (1997). Transients, metastability, and neuronal dynamics. *Neuroimage*. 5, 164–171. doi: 10.1006/nimg.1997.0259

Friston, K. J. (2009). Modalities, modes, and models in functional neuroimaging. *Science* 326, 399–403. doi: 10.1126/science.1174521

Gireesh, E. D., and Plenz, D. (2008). Neuronal avalanches organize as nested theta-and beta/gamma-oscillations during development of cortical layer 2/3. *Proc. Natl. Acad. Sci. U.S.A*. 105, 7576–7581. doi: 10.1073/pnas.0800537105

Golos, M., Jirsa, V., and Daucé, E. (2015). Multistability in large scale models of brain activity. *PLoS Comput Biol*. 11, e1004644. doi: 10.1371/journal.pcbi.1004644

Gratton, C., Nomura, E. M., Pérez, F., and D'Esposito, M. (2012). Focal brain lesions to critical locations cause widespread disruption of the modular organization of the brain. *J. Cogn. Neurosci*. 24, 1275–1285. doi: 10.1162/jocn_a_00222

Haken, H. (1991). "Synergetics – can it help physiology," in *Rhythms in Physiological Systems*, eds H. Haken and H. P. Koepchen (Berlin; Heidelberg: Springer), 21–31.

Herzog, M. H., Kammer, T., and Scharnowski, F. (2016). Time slices: what is the duration of a percept? *PLoS Biol*. 14, e1002433. doi: 10.1371/journal.pbio.1002433

Huys, R., Perdikis, D., and Jirsa, V. K. (2014). Functional architectures and structured flows on manifolds: a dynamical framework for motor behavior. *Psychol. Rev*. 121, 302. doi: 10.1037/a0037014

Joglekar, M. R., Mejias, J. F., Yang, G. R., and Wang, X.-J. (2018). Inter-areal balanced amplification enhances signal propagation in a large-scale circuit model of the primate cortex. *Neuron* 98, 222–234. doi: 10.1016/j.neuron.2018.02.031

Kelso, J. S. (2012). Multistability and metastability: understanding dynamic coordination in the brain. *Philos. Trans. R. Soc. B Biol. Sci*. 367, 906–918. doi: 10.1098/rstb.2011.0351

Kim, H., and Lee, U. (2019). Criticality as a determinant of integrated information ϕ in human brain networks. *Entropy* 21, 981. doi: 10.3390/e21100981

Kinouchi, O., and Copelli, M. (2006). Optimal dynamical range of excitable networks at criticality. *Nat. Phys*. 2, 348–351. doi: 10.1038/nphys289

Koch, C., Massimini, M., Boly, M., and Tononi, G. (2016). Neural correlates of consciousness: progress and problems. *Nat. Rev. Neurosci*. 17, 307–321. doi: 10.1038/nrn.2016.22

Koivisto, M., and Grassini, S. (2016). Neural processing around 200 ms after stimulus-onset correlates with subjective visual awareness. *Neuropsychologia* 84, 235–243. doi: 10.1016/j.neuropsychologia.2016.02.024

Koivisto, M., and Revonsuo, A. (2010). Event-related brain potential correlates of visual awareness. *Neurosci. Biobehav. Rev*. 34, 922–934. doi: 10.1016/j.neubiorev.2009.12.002

Larremore, D. B., Shew, W. L., and Restrepo, J. G. (2011). Predicting criticality and dynamic range in complex networks: effects of topology. *Phys. Rev. Lett*. 106, 058101. doi: 10.1103/PhysRevLett.106.058101

Laurson, L., Illa, X., Santucci, S., Tore Tallakstad, K., Måløy, K. J., and Alava, M. J. (2013). Evolution of the average avalanche shape with the universality class. *Nat. Commun*. 4, 1–6. doi: 10.1038/ncomms3927

Lombardi, F., Herrmann, H. J., Perrone-Capano, C., Plenz, D., and De Arcangelis, L. (2012). Balance between excitation and inhibition controls the temporal organization of neuronal avalanches. *Phys. Rev. Lett*. 108, 228703. doi: 10.1103/PhysRevLett.108.228703

Lombardi, F., Herrmann, H. J., Plenz, D., and De Arcangelis, L. (2016). Temporal correlations in neuronal avalanche occurrence. *Sci. Rep*. 6, 1–12. doi: 10.1038/srep24690

Lombardi, F., Pepić, S., Shriki, O., Tkačik, G., and DeMartino, D. (2021). Quantifying the coexistence of neuronal oscillations and avalanches. *arXiv [Preprint]*. arXiv: 2108.06686. Available online at: https://arxiv.org/pdf/2108.06686.pdf

Lord, L.-D., Stevner, A. B., Deco, G., and Kringelbach, M. L. (2017). Understanding principles of integration and segregation using whole-brain computational connectomics: implications for neuropsychiatric disorders. *Philos. Trans. R. Soc. A Math*., *Phys. Eng. Sci*. 375, 20160283. doi: 10.1098/rsta.2016.0283

Madl, T., Baars, B. J., and Franklin, S. (2011). The timing of the cognitive cycle. *PLoS ONE* 6, e14803. doi: 10.1371/journal.pone.0014803

Mashour, G. A. (2018). The controversial correlates of consciousness. *Science* 360, 493–494. doi: 10.1126/science.aat5616

Mashour, G. A., Roelfsema, P., Changeux, J.-P., and Dehaene, S. (2020). Conscious processing and the global neuronal workspace hypothesis. *Neuron* 105, 776–798. doi: 10.1016/j.neuron.2020.01.026

Mediano, P. A., Rosas, F. E., Farah, J. C., Shanahan, M., Bor, D., and Barrett, A. B. (2022). Integrated information as a common signature of dynamical and information-processing complexity. *Chaos* 32, 013115. doi: 10.1063/5.0063384

Melloni, L., Molina, C., Pena, M., Torres, D., Singer, W., and Rodriguez, E. (2007). Synchronization of neural activity across cortical areas correlates with conscious perception. *J. Neurosci*. 27, 2858–2865. doi: 10.1523/JNEUROSCI.4623-06.2007

Melloni, L., Mudrik, L., Pitts, M., and Koch, C. (2021). Making the hard problem of consciousness easier. *Science* 372, 911–912. doi: 10.1126/science.abj3259

Michel, C. M., and Koenig, T. (2018). Eeg microstates as a tool for studying the temporal dynamics of whole-brain neuronal networks: a review. *Neuroimage* 180, 577–593. doi: 10.1016/j.neuroimage.2017.11.062

Miller, S. R., Yu, S., and Plenz, D. (2019). The scale-invariant, temporal profile of neuronal avalanches in relation to cortical γ-oscillations. *Sci. Rep*. 9, 1–14. doi: 10.1038/s41598-019-52326-y

Nagel, T. (1974). What is it like to be a bat. *Read. Philos. Psychol*. 1, 159–168. doi: 10.2307/2183914

OByrne, J., and Jerbi, K. (2022). How critical is brain criticality? *Trends Neurosci*. doi: 10.1016/j.tins.2022.08.007

Odegaard, B., Knight, R. T., and Lau, H. (2017). Should a few null findings falsify prefrontal theories of conscious perception? *J. Neurosci*. 37, 9593–9602. doi: 10.1523/JNEUROSCI.3217-16.2017

Oizumi, M., Albantakis, L., and Tononi, G. (2014). From the phenomenology to the mechanisms of consciousness: integrated information theory 3.0. *PLoS Comput. Biol*. 10, e1003588. doi: 10.1371/journal.pcbi.1003588

Palva, J. M., Zhigalov, A., Hirvonen, J., Korhonen, O., Linkenkaer-Hansen, K., and Palva, S. (2013). Neuronal long-range temporal correlations and avalanche dynamics are correlated with behavioral scaling laws. *Proc. Natl. Acad. Sci. U.S.A*. 110, 3585–3590. doi: 10.1073/pnas.1216855110

Papadimitriou, C. H., Vempala, S. S., Mitropolsky, D., Collins, M., and Maass, W. (2020). Brain computation by assemblies of neurons. *Proc. Natl. Acad. Sci. U.S.A*. 117, 14464–14472. doi: 10.1073/pnas.2001893117

Papanikolaou, S., Bohn, F., Sommer, R. L., Durin, G., Zapperi, S., and Sethna, J. P. (2011). Universality beyond power laws and the average avalanche shape. *Nat. Phys*. 7, 316–320. doi: 10.1038/nphys1884

Plenz, D., Ribeiro, T. L., Miller, S. R., Kells, P. A., Vakili, A., and Capek, E. L. (2021). Self-organized criticality in the brain. *Front. Phys*. 9, 639389. doi: 10.3389/fphy.2021.639389

Popiel, N. J., Khajehabdollahi, S., Abeyasinghe, P. M., Riganello, F., Nichols, E. S., Owen, A. M., et al. (2020). The emergence of integrated information, complexity, and "consciousness" at criticality. *Entropy* 22, 339. doi: 10.3390/e22030339

Pöppel, E. (1997). A hierarchical model of temporal perception. *Trends Cogn. Sci*. 1, 56–61. doi: 10.1016/S1364-6613(97)01008-5

Poppel, E. (2004). Lost in time: a historical frame, elementary processing units and the 3-second window. *Acta Neurobiol. Exp*. 64, 295–302.

Preti, M. G., Bolton, T. A., and Van De Ville, D. (2017). The dynamic functional connectome: state-of-the-art and perspectives. *Neuroimage* 160, 41–54. doi: 10.1016/j.neuroimage.2016.12.061

Priesemann, V., Valderrama, M., Wibral, M., and Le Van Quyen, M. (2013). Neuronal avalanches differ from wakefulness to deep sleep-evidence from intracranial depth recordings in humans. *PLoS Comput. Biol*. 9, e1002985. doi: 10.1371/journal.pcbi.1002985

Priesemann, V., Wibral, M., Valderrama, M., Pröpper, R., Le Van Quyen, M., Geisel, T., et al. (2014). Spike avalanches *in vivo* suggest a driven, slightly subcritical brain state. *Front. Syst. Neurosci*. 8, 108. doi: 10.3389/fnsys.2014.00108

Rabuffo, G., Fousek, J., Bernard, C., and Jirsa, V. (2021). Neuronal cascades shape whole-brain functional dynamics at rest. *eNeuro* 8, ENEURO.0283-21.2021. doi: 10.1523/ENEURO.0283-21.2021

Roberts, J. A., Gollo, L. L., Abeysuriya, R. G., Roberts, G., Mitchell, P. B., Woolrich, M. W., et al. (2019). Metastable brain waves. *Nat. Commun*. 10, 1–17. doi: 10.1038/s41467-019-08999-0

Rodriguez, E., George, N., Lachaux, J.-P., Martinerie, J., Renault, B., and Varela, F. J. (1999). Perception's shadow: long-distance synchronization of human brain activity. *Nature* 397, 430–433. doi: 10.1038/17120

Salti, M., Bar-Haim, Y., and Lamy, D. (2012). The p3 component of the erp reflects conscious perception, not confidence. *Conscious. Cogn*. 21, 961–968. doi: 10.1016/j.concog.2012.01.012

Sanders, R. D., Tononi, G., Laureys, S., Sleigh, J. W., and Warner, D. S. (2012). Unresponsiveness$\neq$ unconsciousness. *J. Am. Soc. Anesthesiol*. 116, 946–959. doi: 10.1097/ALN.0b013e318249d0a7

Schartner, M., Seth, A., Noirhomme, Q., Boly, M., Bruno, M.-A., Laureys, S., et al. (2015). Complexity of multi-dimensional spontaneous EEG decreases during propofol induced general anaesthesia. *PLoS ONE* 10, e0133532. doi: 10.1371/journal.pone.0133532

Sergent, C., Corazzol, M., Labouret, G., Stockart, F., Wexler, M., King, J.-R., et al. (2021). Bifurcation in brain dynamics reveals a signature of conscious processing independent of report. *Nat. Commun*. 12, 1–19. doi: 10.1038/s41467-021-21393-z

Sethna, J. P., Dahmen, K. A., and Myers, C. R. (2001). Crackling noise. *Nature* 410, 242–250. doi: 10.1038/35065675

Shew, W. L., and Plenz, D. (2013). The functional benefits of criticality in the cortex. *Neuroscientist* 19, 88–100. doi: 10.1177/1073858412445487

Shew, W. L., Yang, H., Yu, S., Roy, R., and Plenz, D. (2011). Information capacity and transmission are maximized in balanced cortical networks with neuronal avalanches. *J. Neurosci*. 31, 55–63. doi: 10.1523/JNEUROSCI.4637-10.2011

Shine, J. M. (2019). Neuromodulatory influences on integration and segregation in the brain. *Trends Cogn. Sci*. 23, 572–583. doi: 10.1016/j.tics.2019.04.002

Shine, J. M., Breakspear, M., Bell, P. T., Ehgoetz Martens, K. A., Shine, R., Koyejo, O., et al. (2019). Human cognition involves the dynamic integration of neural activity and neuromodulatory systems. *Nat. Neurosci*. 22, 289–296. doi: 10.1038/s41593-018-0312-0

Shriki, O., Alstott, J., Carver, F., Holroyd, T., Henson, R. N., Smith, M. L., et al. (2013). Neuronal avalanches in the resting meg of the human brain. *J. Neurosci*. 33, 7079–7090. doi: 10.1523/JNEUROSCI.4286-12.2013

Signorelli, C. M., Szczotka, J., and Prentner, R. (2021). Explanatory profiles of models of consciousness-towards a systematic classification. *Neurosci. Conscious*, 2021, niab021. doi: 10.1093/nc/niab021

Singer, W. (2001). Consciousness and the binding problem. *Ann. N. Y. Acad. Sci*. 929, 123–146. doi: 10.1111/j.1749-6632.2001.tb05712.x

Singhal, I., and Srinivasan, N. (2021). Time and time again: a multi-scale hierarchical framework for time-consciousness and timing of cognition. *Neurosci. Conscious*. 2021, niab020. doi: 10.1093/nc/niab020

Sorrentino, P., Rucco, R., Baselice, F., De Micco, R., Tessitore, A., Hillebrand, A., et al. (2021a). Flexible brain dynamics underpins complex behaviours as observed in parkinsons disease. *Sci. Rep*. 11, 1–12. doi: 10.1038/s41598-021-83425-4

Sorrentino, P., Seguin, C., Rucco, R., Liparoti, M., Lopez, E. T., Bonavita, S., et al. (2021b). The structural connectome constrains fast brain dynamics. *Elife* 10:e67400. doi: 10.7554/eLife.67400

Sporns, O. (2013). Network attributes for segregation and integration in the human brain. *Curr. Opin. Neurobiol*. 23, 162–171. doi: 10.1016/j.conb.2012.11.015

Srinivasan, R., Russell, D. P., Edelman, G. M., and Tononi, G. (1999). Increased synchronization of neuromagnetic responses during conscious perception. *J. Neurosci*. 19, 5435–5448. doi: 10.1523/JNEUROSCI.19-13-05435.1999

Suzuki, M., and Larkum, M. E. (2020). General anesthesia decouples cortical pyramidal neurons. *Cell* 180, 666–676. doi: 10.1016/j.cell.2020.01.024

Tagliazucchi, E. (2017). The signatures of conscious access and its phenomenology are consistent with large-scale brain communication at criticality. *Conscious. Cogn.* 55, 136–147. doi: 10.1016/j.concog.2017.08.008

Tagliazucchi, E., Balenzuela, P., Fraiman, D., and Chialvo, D. R. (2012). Criticality in large-scale brain fmri dynamics unveiled by a novel point process analysis. *Front. Physiol.* 3, 15. doi: 10.3389/fphys.2012.00015

Tognoli, E., and Kelso, J. S. (2014). The metastable brain. *Neuron* 81, 35–48. doi: 10.1016/j.neuron.2013.12.022

Tononi, G., Boly, M., Massimini, M., and Koch, C. (2016). Integrated information theory: from consciousness to its physical substrate. *Nat. Rev. Neurosci.* 17, 450–461. doi: 10.1038/nrn.2016.44

Tononi, G., Edelman, G. M., Libet, B., Koch, C., Gray, J., and De. scarries, L. (1998a). Consciousness and the integration of information in the brain. discussion. *Adv. Neurol.* 77, 245–280.

Tononi, G., Edelman, G. M., and Sporns, O. (1998b). Complexity and coherency: integrating information in the brain. *Trends Cogn. Sci.* 2, 474–484. doi: 10.1016/S1364-6613(98)01259-5

Uhrig, L., Sitt, J. D., Jacob, A., Tasserie, J., Barttfeld, P., Dupont, M., et al. (2018). Resting-state dynamics as a cortical signature of anesthesia in monkeys. *Anesthesiology* 129, 942–958. doi: 10.1097/ALN.0000000000002336

Van Vugt, B., Dagnino, B., Vartak, D., Safaai, H., Panzeri, S., Dehaene, S., et al. (2018). The threshold for conscious report: Signal loss and response bias in visual and frontal cortex. *Science* 360, 537–542. doi: 10.1126/science.aar7186

VanRullen, R. (2016). Perceptual cycles. *Trends Cogn. Sci.* 20, 723–735. doi: 10.1016/j.tics.2016.07.006

VanRullen, R., and Koch, C. (2003). Is perception discrete or continuous? *Trends Cogn. Sci.* 7, 207–213. doi: 10.1016/S1364-6613(03)00095-0

Varela, F. J. (1999). "The specious present: a neurophenomenology of time consciousness," in *Naturalizing Phenomenology: ISSUES in Contemporary Phenomenology and Cognitive Science*, eds J. Petitot, F. J. Varela, B. Pachoud, and J.-M. Roy (Redwood City, CA: Stanford University Press), 266–314.

Walter, N., and Hinterberger, T. (2022). Self-organized criticality as a framework for consciousness: a review study. *Front. Psychol.* 13, 911620. doi: 10.3389/fpsyg.2022.911620

Wanja, W. (2017). "Predictive processing and the phenomenology of time consciousness: a hierarchical extension of Rick Grush's trajectory estimation model," in *Philosophy and Predictive Processing* (Frankfurt am Main: MIND Group), 433–453. doi: 10.25358/openscience-649

Ward, L. M. (2003). Synchronous neural oscillations and cognitive processes. *Trends Cogn. Sci.* 7, 553–559. doi: 10.1016/j.tics.2003.10.012

Werner, G. (2007). Perspectives on the neuroscience of cognition and consciousness. *Biosystems* 87, 82–95. doi: 10.1016/j.biosystems.2006.03.007

Werner, G. (2009). Viewing brain processes as critical state transitions across levels of organization: neural events in cognition and consciousness, and general principles. *Biosystems* 96, 114–119. doi: 10.1016/j.biosystems.2008.11.011

Wiese, W. (2020). The science of consciousness does not need another theory, it needs a minimal unifying model. *Neurosci. Conscious.* 2020, niaa013. doi: 10.1093/nc/niaa013

Zamani Esfahlani, F., Jo, Y., Faskowitz, J., Byrge, L., Kennedy, D. P., Sporns, O., et al. (2020). High-amplitude cofluctuations in cortical activity drive functional connectivity. *Proc. Natl. Acad. Sci. U.S.A.* 117, 28393–28401. doi: 10.1073/pnas.2005531117

frontiers | Frontiers in Psychology

TYPE Review
PUBLISHED 13 February 2023
DOI 10.3389/fpsyg.2022.1028150

Check for updates

OPEN ACCESS

EDITED BY
Antonino Raffone,
Sapienza University of Rome, Italy

REVIEWED BY
Peter J. Marshall,
Temple University,
United States
Luca Simione,
Institute of Cognitive Sciences and Technologies, Italy

*CORRESPONDENCE
Irene Sperandio
✉ irene.sperandio@unitn.it
Federica Meconi
✉ federica.meconi@unitn.it

†These authors have contributed equally to this work and share senior authorship

SPECIALTY SECTION
This article was submitted to Consciousness Research, a section of the journal Frontiers in Psychology

RECEIVED 25 August 2022
ACCEPTED 12 December 2022
PUBLISHED 13 February 2023

CITATION
Dapor C, Sperandio I and Meconi F (2023) Fading boundaries between the physical and the social world: Insights and novel techniques from the intersection of these two fields.
Front. Psychol. 13:1028150.
10.3389/fpsyg.2022.1028150

Fading boundaries between the physical and the social world: Insights and novel techniques from the intersection of these two fields

Cecilia Dapor, Irene Sperandio*† and Federica Meconi*†

Department of Psychology and Cognitive Science, University of Trento, Rovereto, Italy

This review focuses on the subtle interactions between sensory input and social cognition in visual perception. We suggest that body indices, such as gait and posture, can mediate such interactions. Recent trends in cognitive research are trying to overcome approaches that define perception as stimulus-centered and are pointing toward a more embodied agent-dependent perspective. According to this view, perception is a constructive process in which sensory inputs and motivational systems contribute to building an image of the external world. A key notion emerging from new theories on perception is that the body plays a critical role in shaping our perception. Depending on our arm's length, height and capacity of movement, we create our own image of the world based on a continuous compromise between sensory inputs and expected behavior. We use our bodies as natural "rulers" to measure both the physical and the social world around us. We point out the necessity of an integrative approach in cognitive research that takes into account the interplay between social and perceptual dimensions. To this end, we review long-established and novel techniques aimed at measuring bodily states and movements, and their perception, with the assumption that only by combining the study of visual perception and social cognition can we deepen our understanding of both fields.

KEYWORDS

social cognition, visual perception, embodiment, body indexes, posture, empathy, social body

1. General introduction

New insights on how traditional neuroscientific methodologies and instruments can adapt to emerging models of social cognition and perception are long needed. The vast majority of studies on perception so far have mainly concentrated on dissecting bottom-up processes, such as feature categorization and grouping. However, higher cognition seems to be rooted in our sensory representation of the world and to be modeled by our physicality. Therefore, we need research instruments able to tap into the deep embedding of our cognition in bodily processes. After all, the interface between the environment, whether physical or social, and its perception at our end comes down to our body.

For instance, this article focuses on the subtle interactions between sensory input and social cognition in visual perception and proposes the body indices, such as gait and posture, as a critical mediator of such interactions.

We structure our perspective in three sections: the first one starts by exposing different approaches in perception studies and by highlighting the deep interplay between top-down and perceptual processes; the second one focuses on embodied sensorial aspects of social cognition and its reliance on visuospatial behavior mechanisms; the third one presents a selection of studies in which bodily states (physiological, postural, and kinematical) are analyzed in their role of conveying social information to the viewer.

The key idea behind this work is that sociality is a part of – and not apart from – our biological being: we have preconscious dispositions toward perceiving social cues, namely other people's bodies and their expressions, and the way we perceive the (social) environment is shaped by our own body and its expressions. Accumulating evidence from different lines of research is supporting this notion, revealing that our perceptual experience is not a mere copy of the external appearances, but rather the result of a compromise between physical environment's layouts and people's social needs and expectations (e.g., Balcetis and Dunning, 2009; Firestone and Scholl, 2016; Niedenthal and Wood, 2019; Sato et al., 2019; Valenti and Firestone, 2019; Sun et al., 2021).

Building upon compelling evidence showing an overlap between spatial and social behavior in the perception of distances (Henderson et al., 2006; Bar-Anan et al., 2007; Yamakawa et al., 2009; Xiao and Bavel, 2012; Parkinson and Wheatley, 2013; Takahashi et al., 2013; Hamilton et al., 2014; Schiano Lomoriello et al., 2018; Sato et al., 2019; Kroczek et al., 2020), and on studies on the preconscious and automatic processing of social visual stimuli (Morris et al., 2001; Pegna et al., 2005; Tamietto and de Gelder, 2008; de Gelder, 2009; de Gelder et al., 2010; Zhou et al., 2019), we claim that in the process of understanding others, social top-down and visual bottom-up processes join their forces to prompt an adaptive response to the social stimulus. Moreover, we will show that the exploration of our physical and social surroundings is self-centered, or, more precisely, body-centered. In other words, our body shape and appearance strongly affect the way we move and conduct ourselves toward physical and social entities (Yee et al., 2009; Linkenauger et al., 2010; van der Hoort et al., 2011), revealing a fundamental role of our body in shaping our perception and our sociality.

In short, the main claim of this review is that our body is a critical mediator of the interplay between social cognition and visual perception; the body is indeed our interface at the basis of our interaction with the environment. We can trace the effects of this mediation even in the rescaling effects that it exerts on our sight: understanding the deep influence of our bodies and bodily processes on our visual perception is a crucial step to frame also its role in shaping our attitude toward and interpretation of social relationships. We observe other people's postures and movements in order to understand their emotions or intentions, and the effective recognition of the others' psychological state is accompanied by the instinctive simulation of their expressions in our own body. Theories of embodied cognition (Goldstone and Barsalou, 1998; Niedenthal et al., 2005; Proffitt, 2006; Gallese, 2009) claim that we use sensorimotor models originating from our proprioception and body schema (see definition in Box 2) to infer the psychological state of others. However, the literature often lacks the specification of the exact bodily indexes observed and simulated during social processes. We want to help fill this gap by presenting and promoting novel techniques that integrate the study of social cognition with physiological and visual measures, aiming at incentivizing integrative and multimodal approaches. To sum up, we strongly encourage the re-integration of the whole body in the study of social behavior and visual perception, following Nakayama's claim that "vision science is going social" (Nakayama, 2011).

2. Reconstructing visual perception

The concept of perception has evolved throughout history. The last century has seen different theories succeeding one another in an attempt of giving a systematic explanation of how humans perceive the world surrounding them. Since the first definitions of perceptual processes, a differentiation between early sensory systems and higher cognitive processes has emerged. Although a reciprocal influence between these different levels of processing is well-established in the literature, an unsolved issue concerns the stage at which these two levels start interacting. For instance, traditional views assume a hierarchical structure in which the external input is initially processed by sensory cortices and it is only later ri-elaborated and interpreted by higher cognition. In contrast, alternative approaches (e.g., New Look, enactivism, ecological psychology, 4E approaches) argue for an early intervention of motivational systems, which modulate our perception by increasing the saliency of some objects on the basis of our current needs or desires. Within this other theoretical framework, perceptual processes have a constructive nature for which external information is gathered and modeled for the demands of our inner states (Balcetis and Dunning, 2006).

2.1. New look and other perspectives in perception: Visual perception as a constructive process

According to the New Look theory, perception is far from being a neutral and objective representation of the external world. Instead, what we see appears to be the result of a compromise between autochthonous and behavioral determinants (Bruner and

10.3389/fpsyg.2022.1028150

Goodman, 1947). While *autochthonous* refers to the electrochemical signals generated by the sensory end organs, *behavioral* includes learning and motivation, social needs and attitudes, basic physiological needs, such as hunger and thirst. These two determinants build a perceptual hypothesis that undergoes a selection process driven by our needs or requirements. Objects that comply with the selective criteria become "more vivid, have greater clarity or greater brightness or greater apparent size" (Bruner and Goodman, 1947). In a classic experiment, Bruner and Goodman (1947) showed that a physical entity, like a metal disk, invested of social value (e.g., a coin) is perceived as having a bigger shape compared to the same form without any utility (i.e., a simple metal disk). The authors postulated that the greater the need for a socially valued object, the more marked the reshaping of the perceived entity would be. In agreement with this hypothesis, they demonstrated that poorer children were more likely to perceive the coins as bigger, compared to an age-matched group of more wealthy individuals (Bruner and Goodman, 1947).

Despite initial criticism, this theory has recently regained attention thanks to scholars who have developed rigorous methodological approaches and demonstrated how semantic knowledge (e.g., stereotypes) can shape - at a very early stage – the way we process sensory inputs (Otten et al., 2017).

Two additional influential theories that date back to the second half of the last century and originated in opposition to cognitivism, are the enactivism and ecological psychology. The latter was pioneered by Gibson (2014) and has its foundations in his perspective on visual perception. The most relevant point of this framework is the concept of affordances (see Box 1), a notion that has been exploited in different fields and transposed from the physical environment of inanimate objects to the social affordances offered by another person. Another fundamental aspect of ecological psychology is the reintroduction of the body as a reference point for our perceptions, and the concept of perception as an act of information pickup. Gibson revolutionized the whole field of visual perception, introducing new models based on the way that our visual sensory system is built on the idiosyncrasies of our body rather than the external stimulus.

While one can easily consider ecological psychology as a new theory of perception, the redefinition of the organism-environment relationship imposed by this theory has important implications for cognition (Popova and Rączaszek-Leonardi, 2020). A more radical elaboration of the centrality of our body in cognitive processes has been put forward by the enactivist approach, largely founded on the ideas of Varela et al. (1991). Similar to ecological psychology, this theory strongly highlights the importance of the relationship between agent and environment, with the assumption that cognition and perception emerge together and are so closely connected that they cannot be considered separately. The objectivity of perception is deconstructed in favor of the inevitability of the subjective state in the global landscape formed by the individual immersed in its surroundings (Fuchs, 2020; Heft, 2020; de Pinedo García, 2020; Popova and Rączaszek-Leonardi, 2020; Read and Szokolszky, 2020).

This account of perception conforms also to the more and more popular theory of predictive coding (PC), which has its origin in the model of visual perception formulated by Rao and Ballard (1999). The authors described visual perception not only as a feedforward loop from lower-to higher-order visual cortical areas, but also as a cycle of prediction and error-correction. The new emerging ideas that followed this initial assertion, are based on the inferential nature of the brain (Friston, 2018). In a nutshell, the PC framework states that our neural networks constantly predict sensorial aspects of the environment based on the statistical regularities of the natural world. This prediction generates a perceptive model which is confronted with the incoming sensory inputs and corrected in case of mismatch (Rao and Ballard, 1999; Friston and Kiebel, 2009). One can easily see the parallelism with Bayesian modeling in statistics, for which an hypothetical or aprioristic model of a certain phenomenon is continuously updated on the basis of the newly collected data. What the system, in this case the brain, should focus its energy on is to minimize the prediction error (Friston and Kiebel, 2009; Friston, 2018). This inferential view of our cognition and sensorium, described also as Baesyan brain hypothesis, has gained growing consensus and has become one of the most dominant models in cognitive neuroscience. The PC approach has been applied in multiple fields of science. A relevant example is its application to explain interoceptive awareness or accuracy (Seth et al., 2012; Ainley et al., 2016). In the study of Ainley et al. (2016), the sensitivity for internal bodily information is explained as the ability to adapt a "prior" representation of the body state to the effective interoceptive sensation, by continuously adjusting and minimizing the prediction errors. If this predictive dynamics holds true for interoceptive inputs as well and not only for what we expect to perceive from the external environment, then the model can be extended to partially explain our emotional responses (Seth, 2013; Barrett and Satpute, 2019). In fact, emotions can be seen as predictions of reactions to an event based on past experiences that generated models of behavior. They can represent the active inference of how to respond to a particular situation, and be created from memory with the possibility of future adjustments. In a more articulated manner, Seth (2013) integrates the interoceptive predictive coding hypothesis in a coherent understanding of emotional content as an "active top-down inference of the causes of interoceptive signals." Interoception is no longer considered as a one-directional collection of bodily sensations from peripheral receptors to the central processor, but rather as a process shaped by top-down inferences and influences along with bottom-up error-correction processes. In summary, the predictive coding account of emotions defines them as active inferences on the causes of physiological changes, both at internal and external representational levels (Seth, 2013; Barrett and Satpute, 2019).

Again, the "strange inversion" brought about by this theoretical account goes in the same direction as the other frameworks described so far, which point out the inferential, constructive, and active functioning of the perceiving organism, BOX 1 Definition of affordances.

The term was introduced by James Gibson, within its ecological approach to the study of vision, and indicates the totality of actions that an object allows to perform. The affordances of an object can be defined as the potential actions elicited by the view of that object. For example, seeing an apple suggests the actions of eating it, grabbing it, or moving it. A chair "affords" being sat on, and so forth.

whilst leaving behind the idea of the brain "as a glorious stimulus–response link" (Friston, 2018). Following this discussion, it is probable that the "naive realism," supported by traditional approaches, whereby what we see, smell and feel is a faithful representation of what is "out there," will be progressively replaced by a view of our visual system as grounded in action capabilities and social influences (Proffitt and Linkenauger, 2013; Proffitt and Baer, 2020). The current review aims at pushing cognitive research into this direction, highlighting the deep influences that our inner states and social needs exert on perceptual processing. In the ensuing sections, we will argue that the visual salience of an object is given by its affordances (see definition in Box 1) to satisfy our physiological or social needs. In other words, we project on the objects surrounding us different degrees of desiderability – based on our bodily and psychological states - and perceive them accordingly, inevitably bonding our perception to our action possibilities.

2.2. Our metabolic energies influence visual perception

Our action possibilities are delimited by the resources we dispose of and the costs of the actions we want to undertake. Using Proffitt's words, "survival for any organism, including people, is a matter of resource management" (Gross and Proffitt, 2013). Our brain calculates costs and opportunities associated with every movement, and this constant evaluation is carried out automatically and unconsciously, in that, if we had to be aware of it, our executive functions would overload (Proffitt, 2006; Gross and Proffitt, 2013).

First and foremost, our ability to perform an action depends on our body characteristics: our body size determines what we can reach as well as what we can see (Linkenauger and Proffitt, 2008; Sugovic et al., 2016). The effect of our body mass on perceived distance was assessed in an experiment of Sugovic et al. (2016) in which normal weight, overweight, and obese individuals were asked to estimate a same distance and report their beliefs concerning their body size. They found that perceived distance was mainly affected by physical body weight and that this effect was independent from personal beliefs. Specifically, they found that the heavier the person, the greater the estimation of distance (Sugovic et al., 2016). Also, the physiological state of our body plays a major role in what we can perform and how. Being tired, or out of training, or carrying a weight, are all factors that can diminish our potential to perform actions (Proffitt, 2006; Linkenauger and Proffitt, 2008; Proffitt and Baer, 2020). A reduction of our action possibilities translates into an adjustment of the environmental perception. In fact, it has been shown that perceived distances increase when our energies are scarce and are instead reduced when we are trained and performative (Zadra et al., 2016; Proffitt and Baer, 2020). In the same fashion, Bhalla and Proffitt (1999) showed that the steepness of a hill was overestimated by participants who were asked to carry a backpack, fatigued runners or those in low physical and health conditions when compared to participants in their full forces (e.g., not carrying a backpack or fit and in good health). A further example on how sugar intake and fitness level affect visual estimates was provided by Zadra et al. (2016). The authors asked participants to judge distances after physical exercise. Prior to the physical activity, half of the participants received a carbohydrate supplement, whereas the other half received a placebo. They observed that those who received the energizer rated the extent to be shorter compared to the placebo group. They also found that perceived distance correlated with other measures of fitness, such as blood glucose, heart rate (HR), and caloric expenditure under physical fatigue, further confirming the influence of bioenergetic resources on perceptual processing (Zadra et al., 2016). Interestingly, Changizi and Hall (2001) demonstrated that thirsty people perceive ambiguous visual stimuli as more transparent than non-thirsty subjects, and this may be due to the implicit association of transparency with water.

2.3. Tools extend our action possibilities modifying the perception of our surroundings

Beside dimensions and fitness of our body, another factor that can determine our potential to perform actions is the use of tools. In fact, tools can extend our reach to extrapersonal space, including farther objects in our area of manipulability. Objects within our reach are automatically perceived as candidates for potential actions and this implies different perceptual processings (for a review, see Brockmole et al., 2013). Indeed, they are visually scanned in a more attentive way and with a detail-oriented processing style in order to enable appropriate action responses compared to objects we cannot touch or immediately interact with (Brockmole et al., 2013). Hence, holding a tool that increases the area of our possible interaction with the surroundings immediately modifies the visual perception of what would have been beyond our reach. For example, when patients with hemispatial neglect (a neuropsychological condition characterized by reduced awareness of visual stimuli in one side of the visual field, not accompanied by sensorial deficit) limited to the near space were provided with a stick and asked to perform a line bisection task in the far space, the neglect expanded to include the area reachable with the tool (Berti and Frassinetti, 2000). This demonstrates that the artificial extension of our reach remodels peripersonal space and the perception of far and near objects, suggesting once more that our chance of interaction with the environment has deep effects on how we perceive it.

2.4. Social baseline theory: Social resources can directly alter our visual perceptions

Not only metabolic energies weigh on the capacity to undertake any action, but also they influence our social relationships. Supportive social networks allow distributing the efforts of any endeavor and protect the individual from potential dangers. It has even been argued that receiving help from other humans is a matter of survival and that the greatest human strength are other humans (Oishi et al., 2013). Indeed, we are born and raised within a social environment that provides for our basic needs until we can take care of ourselves. Even then, we are for the rest of our lives embedded in social networks (Gross and Proffitt, 2013). Interestingly, similarly to what has been observed for metabolic resources, it has been demonstrated that social resources, too, influence our perception of the environment.

The Social Baseline Theory (SBT), formulated by Beckes and Coan (2011), describes interindividual differences in reacting to social support. According to this theoretical account, social support is considered as a default precondition of our actions and determines the baseline from which we calculate the amount of energy available (Beckes and Coan, 2011; Cole et al., 2013; Gross and Proffitt, 2013). Nevertheless, investigations on the role of social resources in perception have been largely neglected in the literature, especially if compared to the wealth of evidence on the role of physiological states (Gross and Proffitt, 2013). However, accumulating evidence indicates that the presence of a significant other acts as an empowering factor when facing difficulties or specific tasks. For example, Doerrfeld et al. (2012) observed that when participants were asked to lift a box they tended to judge its weight as lighter if they knew they would receive help, compared to when they knew they would lift the box without any help. Schnall et al. (2008) replicated Proffitt's study on slant perception where slant estimates varied as a function of the physical and health conditions of the observer (see subsection 2.2), but this time the social support factor was also considered. Participants could either be accompanied by a friend or imagine the presence of another person (friendly or not). In both cases, social support decreased the perception of the steepness of the hill. Other studies have tried to shed light on the ways in which social inclusion or exclusion impact our sensory systems. These studies revealed a wide range of effects on visual perception. For instance, the feeling of being understood is another factor that appears to influence distance and steepness perception. In an experiment by Oishi et al. (2013), participants were judged by an evaluator with a few adjectives chosen from a list, from which the same participants had previously picked a few words to describe themselves. In the understanding condition, participants received the same judgments they also had chosen, while in the misunderstanding condition the evaluator judged the participants with words that largely differed from the participants' self-assessment. After the evaluation session, participants performed a slant estimation task similar to the one used in Bhalla and Proffitt's study (1999). Results showed that feeling understood decreased the slope estimates compared to the misunderstanding condition (Oishi et al., 2013).

To conclude, we can claim that vision assists our possibility of action by modulating perceived size and distance of objects. Our potential to perform actions, in turn, is determined by our dispositional bioenergetic and social resources. Therefore, we see the environment according to the dispositions of our bodily states and of our social network.

Table 1 provides a summary of the studies cited in this section.

3. Perceiving the physical and the social world

After having grounded our discussion in a constructive and embodied perspective of visual perception, we can now focus on its integration in models of social cognition. We will start by drawing a parallel between the research approaches used in visual perception and those used in social cognition. In both cases, we will conclude that adopting a multimodal integrative perspective can better represent the intertwined and complex nature of these processes. We will then present the embodied account of social cognition, which we believe to be the most accurate explanation of how we understand one another. Finally, we will provide empirical evidence for a common mechanism for mapping social and physical distances, a further confirmation of the tight link between social cognition and visuospatial perception mediated by the processing of body indices.

3.1. A parallel between approaches in perception and social cognition

The issue of how we represent objects and events in our mind is and has always been a central theme in cognitive sciences and for a long time these representations have been described as symbolic and amodal. In the field of perception, for example, the main assumption was that we construct abstract representations of the external inputs through mechanisms of feature extrapolation and categorization (Lindblom, 2020). Such a view takes inspiration from Fodor's modularity, according to which encapsulated perceptual modules in the brain transmit the sensory information to higher processing levels that manipulate them in the form of symbolic representation (Fodor, 1983). The same form of representation - amodal and disembodied - has also been used in the study of social cognition. According to traditional accounts, people process social information by means of categories, schemata, feature lists, semantic networks, and so forth (Landau et al., 2010). However, despite their clarity and linearity, these

TABLE 1 Summary table of the studies cited in section 2.

Study	Sample	Stimuli	Design	Results
Sun et al. (2021)	*N* = 18 (13 females, mean age = 22.9)	Target rectangles placed of the top of different stimuli	The different stimuli were selected to induce positive (squirrel), negative (rats), or neutral (wooden blocks) emotions. Participants were asked to rate distances from and size of the targets	Participants perceived the target object on top of the toy rats (induced negative stimulus) significantly closer and larger than the same target object resting on top of the toy squirrels (induced positive stimulus) suggesting that there is a perceptual bias even within reachable personal space
Balcetis and Dunning (2009)	*Experiment 1:* *N* = 63 *Experiment 2:* *N* = 43	Ambiguous figures *Experiment 1:* Number-letter figure, B / 13. *Experiment 2*: Seal-horse figure	Taste-testing of a desirable and a non-desirable beverage. The choice of the beverage/food was based on seeing one of the two possible interpretations of the ambiguous stimulus	Participants' desire to obtain the desirable reward influenced their interpretation of the ambiguous figure
Balcetis and Dunning (2009)	*Experiment 1*: *N* = 90 *Experiment 2a*: *N* = 123 *Experiment 2b:* *N* = 89	Desirable objects *Experiment 1*: A bottle of water *Experiment 2a*: A 100 dollar bill *Experiment 2b:* A survey containing self-relevant feedback	*Experiment 1.* Participants were given either a salty snack or a glass of water and then were asked to estimate the distance between them and a bottle of water *Experiment 2a*. Participants were offered the chance of winning a 100 dollar bill in a card game or only a candy and then asked to estimate the distance from the bill *Experiment 2b.* A survey on sense of humour was ostensibly graded by the experimenter and participants were given either positive or negative feedback and the survey was then placed away from the person and asked to judge how far away	Perceptions of distance depend in part on the desirability of the perceived object—which depends, in turn, on its capacity to satisfy a visceral or intrapsychic need
Zadra et al. (2016)	*N* = 8 (3 females, mean age 26.38)	Walkable distances	A host of physiological measures were recorded as participants engaged in exercise on 2 occasions: once while provided with a carbohydrate supplement and once with a placebo. Distance estimates were made before and after exercise on both occasions	The carbohydrate manipulation caused decreased distance estimates relative to the placebo condition. Individual differences in physiological measures that are associated with physical fitness predicted distance estimates both before and after the experimental manipulations
Bruner and Goodman (1947)	*N* = 30 (mean age = 10)	Disk of different sizes. The disks could either be neutral metal items or coins of different value	The children were asked to estimate the size of the metal disk/coin by adjusting the diameter of a circle of light projected on a screen	Coin size was overestimated, while the neutral disk size estimates were closer to reality. The higher the coin value, the bigger the overestimation. Poorer children overestimated more than more wealthy children the size of the coins
Changizi and Hall (2001)	*N* = 74	Visual stimuli "definitely," "ambiguously," or "definitely not" transparent	Participants had to judge the stimulus as "transparent" or "not transparent" in two conditions: after having eaten a bag of chips (thirsty group) or after having drunk water (non-thirsty group)	The thirsty group showed a greater inclination to judge as transparent the ambiguous stimuli
Bhalla and Proffitt (1999)	*Experiment 1.* *N* = 130 (65 females) *Experiment 2.* *N* = 40 (20 females) *Experiment 3.* *N* = 74 (35 females)	Slant of a hill	Participants were asked to judge how steep the hill was in three ways: verbally, visually, and haptically *Experiment 1.* Participants performed the task in two conditions: wearing a backpack and without it. *Experiment 2.* Participants had to give slant estimates of a hill before a long run (45 to 75 min) and of another hill after the run *Experiment 3.* Different fitness measures were taken from the participants (heart rate in different conditions and body mass index) after or before asking them to judge the slant of the hill	All the experiments demonstrated an effect of load, fatigue, or fitness on the slant estimates, but only for verbal and visual assessments *Experiment 1.* Wearing the backpack influenced the slant estimate making the hill look steeper. *Experiment 2.* The hill estimates were higher after the exhausting run *Experiment 3.* A higher level of fitness corresponded to a lesser overestimation of the hill's slant
Sugovic et al. (2016)	*N* = 66 (30 female, mean age = 24.4)	A cone presented on the sidewalk (4 target distances)	After making four distance estimates, participants completed a survey. Along with demographic questions, the survey asked them to indicate their height, weight, and an evaluative measure of body size	A person's body weight influenced perceived distance: Those who weighed more than others perceived distances to be farther

(Continued)

TABLE 1 (Continued)

Study	Sample	Stimuli	Design	Results
Studies involving social manipulation				
Doerrfeld et al. (2012)	N = 43 (mean age = 22.5)	Box weight	Participants were asked to judge the weight of a box filled with potatoes in two conditions: alone or with someone else	Participants in the joint condition judged the weight as lighter than in the solo condition
Harber et al. (2011)	*Experiment 1.* N = 107 (63% female; mean age = 20.8)	*Experiment 1.* Distance estimates at 3 measurements points	*Experiment 1.* Stimuli were either a live tarantula or a cat toy. Participants were primed with positive, negative or neutral self-worth conditions	Resources moderate the perception of physical threats (dangerous animals, hazardous heights), and are not limited to the implicit calculus of metabolic costs (i.e., how much physical effort a situation might demand, relative to one's physical resources)
	Experiment 2. N = 91 (64.9% female; mean age = 20.18)	*Experiment 2.* Height estimates from the fourth floor of a building	*Experiment 2.* Participants were left free to hold on the handrail or were denied this possibility by tying their hands behind their back. Self-esteem was manipulated	
Cole et al. (2013)	N = 48 (100% females)	A male confederate	Participants were shown a video of the male confederate in which he would appear threatening, disgusting or neutral. Then they were asked for distance estimates	Experimentally induced social signals of threat (but not disgust) led to perceived proximity
Oishi et al. (2013)	N = 202 (112 female; undergraduate students)	Pain endurance (hand in icy water), slant perception, distance perception	Feelings of understanding or misunderstanding were induced in the participants by judging them with a list of adjectives (positive or negative)	Participants in the understanding condition were able to put their hands in icy water for a longer period of time, perceived the target locations to be closer, and perceived the same hill to be less steep than those in the misunderstanding condition
Schnall et al. (2008)	N = 34 (19 female; mean age = 19.94 years)	Slant of a hill	Participants judged the hill slant verbally, visually, and haptically in different conditions: alone or with a friend	Participants with a friend, compared to those alone, saw the hill as less steep. The longer participants knew their friends, the less steep they estimated the hill to be, on both the verbal and visual measures

theories do not account for the multimodal nature of perceptual and social experiences, in which high-and low-level cognitive processes strongly interact (Zaki and Ochsner, 2012). More recently, models in which perception and cognition behave as coupled systems are gaining new ground. In the same way, alternative paradigms of social cognition stemming from theories of embodied cognition (Goldstone and Barsalou, 1998; Niedenthal et al., 2005; Gallese, 2009) are challenging the idea of amodal representations of the social information. We endorse the adoption of multimodal and integrative models of both perception and social cognition, confident that without acknowledging a common basis for perceptual and conceptual processing of physical and social events, our understanding of the brain and the mind would remain incomplete.

3.2. An embodied account of social cognition

Empathy is the ability to understand others' inner state by explicitly inferring it from available contextual information or by internally simulating it (Zaki and Ochsner, 2012; Sessa et al., 2014; Meconi et al., 2018). According to the embodied account of social cognition, we understand other people's mental state by reproducing it in ourselves (e.g., Niedenthal et al., 2005). This is achieved by internally mimicking the same sensorimotor patterns observed in others, which recall specific psychological states we experienced in association with that physical expression (Gallese, 2007, 2009). This has already been shown in a study by Duclos et al. (1989) where participants were asked to mimic some negative emotion expressions (fear, sadness, anger) by contracting specific muscles. In a first experiment, the expressions were limited to the face, while a second testing involved a full body simulation. Participants were convinced that the study regarded brain lateralization and that the muscles' contraction was a conflicting task, the function of which was to overload the cognitive system. Finally, they had to report their feelings throughout the experiments, choosing among different emotions and rating their intensity. Although participants were naïve to the aims of the experiment, they reported higher intensity for the emotions they were mimicking in that moment, both for facial and full body expressions, giving strength to the idea that the activation of specific sensorimotor schemas elicits those embodied feelings.

The discovery of the mirror neurons is a fundamental step at the basis of embodied social cognition because mirror neurons are considered as one route to the development of our ability to understand others' actions. Human mirror neurons seem to be widely spread across the brain with peaks of concentration in the premotor and somatosensory cortices (Gallese, 2007; Fabbri-Destro and Rizzolatti, 2008; Bastiaansen et al., 2009; Gallese, 2009; Keysers et al., 2010; Mukamel et al., 2010). The peculiarity of these neurons lies in the fact that they fire not only when we perform a specific action, but also when we see that same action performed by someone else. Traditionally, research on human mirror neurons adopted fMRI investigations to identify which areas become more active during the observation of another person, and this has allowed the mapping of the neural circuitries that exhibit mirroring properties. Recently, Paradiso et al. (2021) reviewed the scientific production across species to reveal which brain areas are involved in empathic reaction. Empathy is the ability to resonate with the others' inner state and to explicitly understand it (often referred to as affective and cognitive empathy, respectively). Numerous studies on animal models showed converging results: the anterior cingulate cortex and the amygdala resulted as the main areas involved in empathy-related phenomena. The same authors also reviewed the literature on the role of analgesics in modulating prosocial behavior which shows that reducing pain perception hinders the ability to empathize with the pain of others. In fact, the most recent trends in research on empathy focus on the mechanisms of empathy for pain (Rütgen et al., 2015, 2018; Lamm et al., 2019). This new research direction sought to provide mechanistic explanations to simulation models, by selectively disrupting specific subprocesses - nociception in this case - with different techniques (e.g., tDCS, analgesics) to verify their involvement in cognitive processes, like empathy for pain (Bonini et al., 2022; Maggio et al., 2022). The empathic experience of others' pain has been widely examined by Rütgen and Lamm, who conducted several studies on the role of our own nociception in the ability to recognize and understand others' pain. In an fMRI experiment of 2015, the researchers manipulated participants' nociception (i.e., the encoding of noxious stimuli) by means of placebo analgesia induced in half of the participants. FMRI data was collected, while a painful electrical stimulation was delivered either to the participant or to another person present in the scanner room. Results showed reduced activation of the anterior insular and midcingulate cortex, areas typically involved in empathic responses for pain, in the group of participants in which placebo analgesia was administered compared to the control group in which participants did not receive any treatment. Along with other evidence (Rütgen et al., 2015, 2018), these findings are in line with those studies showing that incidental (Forkmann et al., 2015) and voluntary (Fairhurst et al., 2012) reinstatement of an autobiographical pain, involves the partial recruitment of the brain areas that encoded nociceptive stimuli at the time of memory formation. Indeed, memories of autobiographical physical pain augment participants' cognitive empathy for other individuals depicted in similar physically painful situations (Meconi et al., 2021).

Neuromodulation and lesion studies are also suited to pinpoint the networks underlying mechanisms of embodied social cognition. Such an example is the experiment of Lenzoni et al. (2020), which demonstrated a causal relationship between body expression and emotion recognition. Using a matching task of faces and bodies, the authors measured social abilities in patients with myotonic dystrophy, a neuromuscular disease which induces strong sensorimotor limitations. The clinical population performed significantly worse than the group of healthy controls, demonstrating a causal role of visuomotor abilities in emotion recognition. In a review by Keysers et al. (2018), neuromodulation and lesion studies were presented as evidence for the fundamental role of the primary somatosensory cortex and mirror neurons network (parieto-premotor areas) in understanding and predicting the actions of others, which is in turn connected with the ability of recognizing their emotions.

BOX 2 Definitions of proprioception and body schema.

Proprioception is the capacity of perceiving and detecting the position of our own body in space, as well as the state of activation of our muscles, without the support of our sight. The collection of combined signals from sensory receptors in the muscle, skin, and joints, allows us to be aware of our limbs position and movements, and it is - in this sense - a fundamental aspect of motor control. In fact, proprioceptive information is integrated in our body schema, which combines the peripheral inputs with central (brain) processes in order to lead the execution of any action or movement. Body schema has been defined as the body representation for action. Indeed, it can extend to incorporate any tool we are holding allowing the sophistication of human tool-use abilities

The relevance of bodily states in our social judgments is also rooted in our language. For example, feelings of affection and love are usually described as warm, such as the experience of a hug, while loneliness and social distance are typically associated with cold attributes (e.g., "giving someone a cold shower," "cold-hearted"). Ijzerman and Semin (2009) provided evidence for this deep interdependence between language, perception and social behavior. The authors prompted different temperature conditions by asking their participants to rate the social proximity they felt with a known person of their choice while they were holding cold or hot beverages. As it turned out, the warm condition was associated with greater social proximity, compared to the condition of holding a cold beverage.

Taken as a whole, these findings suggest that bodily experiences can play a preconscious and automatic role in shaping explicit awareness and in leading our interaction with the world. We can even state that without embodying our own and others' psychological states, we are denied the possibility of understanding them. Such a conclusion leads again to the necessity of adopting an integrative approach for studying both perceptual and cognitive mechanisms. In the next subsection, we provide more evidence for the reliance of cognitive processes on perceptual ones, by showing that we recruit the same neural networks dedicated to visuospatial representations of distances to represent different degrees of social proximity.

3.3. When the social meets the spatial: Interpersonal distances

A large body of literature highlights that we use overlapping systems for assessing social proximity and physical distances. For instance, Bar-Anan et al. (2007) used a Stroop-like task in which words indicating close or distant social affiliation ("us" or "enemy") were positioned in closer or farther perspectives. Participants had to indicate if the item's position on the screen was proximal or distal, independently from the meaning of the word. It resulted that words were classified faster when the psychological and the spatial distances were matching, compared to when the two types of distance were incongruent. For example, when the word "us" was written in a close-up position in the scenario, the response time was shorter compared to the condition in which the same word (indicating social proximity) was positioned in the background of the scenario. The authors interpreted this finding in terms of a common mechanism for the processing of spatial and psychological distances (see definition in Box 3), which would explain the slower response in the incoherent condition due to the activation of incongruent representations on the same neural path.

Another important line of research supporting this view is the one that investigates the interpersonal distance in social interaction. It is commonly known that we adjust our position in relation to our intimacy with the people around us (Hall, 1963; Hall et al., 1968; Lenzoni et al., 2020). This effect has been named and described in multiple ways. For instance, Teneggi et al. (2013) defined peripersonal space as a multisensory-motor interface between body and environment and showed that its shrinkage or extension depended on the presence and interaction with others. In this vein, Serino (2019) extensively reviewed the literature on peripersonal space, highlighting the stretchable nature of this multisensorial space and its role in mediating body-environment interactions. Furthermore, the author claimed that this physiological construct has the psychological consequence of defining the boundaries between ourselves and the external world, enabling bodily self-location and consciousness (Serino et al., 2013; Blanke et al., 2015; Noel et al., 2018). It is also suggested that peripersonal space plays an important role in the body–body interactions with other people.

To study precisely this body–body dynamics, Kroczek et al. (2020) used Virtual Reality (VR) to manipulate interpersonal distance in social interactions. Participants had to interact with one of two virtual agents represented in the VR scenario. They were instructed to approach them and start interacting as soon as the agent would look up at them. The authors manipulated the distance of interaction by delaying the moment in which the virtual agent would notice the participant. They found that the closer the participant had to get to the virtual agent in order to be noticed, the "more arousing, less pleasant, and less natural" the interaction was felt. Perception of close distances was also accompanied with increased levels of skin conductance. These results are consistent with the principles of Proxemics, whereby personal space is organized in concentric areas that determine the level of ease we feel being close to another person, i.e., we can empathize with them, which is based on our level of intimacy with that person (Schiano Lomoriello et al., 2018).

Proxemics is not the only discipline that has dealt with concepts of personal distances. Construal Level Theory has also attempted to explain the relationship between social, physical, and temporal distances in terms of psychological dimensions. What is meant by psychological dimension is the level of specificity or abstraction, by which information is represented, that goes from a low-level (incidental and specific) representation of events near us to a high-level (general and prototypical) representation of farther events (Henderson et al., 2006; Trope and Liberman, 2010). The possibility of a shared mechanism for the perception of these different dimensions of distances has been corroborated by fMRI studies, showing activation of the same neural network during the processing of social and physical distances. For example, in an fMRI study, Yamakawa et al. (2009) investigated the role of the parietal cortex in analytic representations of egocentric mapping, which is employed for processing both physical and social relationships. The authors asked participants to perform two tasks. In the first task, participants had to evaluate their physical distance to neutral objects displayed on a screen. In the second task, participants were shown with two faces and had to choose the one with which they felt more compatible. Hemodynamic response was collected during both tasks and revealed a common activation in the parietal cortex. The social distance task was also linked with the activation of extended regions dedicated to social cognition processes, such as the fusiform gyri, the bilateral medial frontal cortices, the inferior frontal cortices, the insular cortices, the left basal ganglia, and the amygdala. Nevertheless, the overlap in the parietal cortex seems to confirm a common neural substrate for the evaluation of spatial and social distances, and indicates that this area is part of the network dedicated to the processing of social stimuli.

It has been argued that the parietal cortex organizes complex social information in a self-referred map of social distances, guiding our spatial behavior toward others (Abraham et al., 2008; Yamazaki et al., 2009; Parkinson and Wheatley, 2013). This supports the idea that visuospatial perception and social cognition are interconnected processes, subserved by a common substrate in the brain. The reciprocal influence of these two kinds of distances is becoming more and more evident in the literature. For example, Schiano Lomoriello et al. (2018) have demonstrated that when people are physically distant from us, we are less prone to empathize with them. In other words, the feeling of social distance or proximity is modulated by the physical distance between us and the other person. The effects are visible also the other way around, in that social inferences (e.g., categorization and stereotyping) can tweak our perception of the physical world, as demonstrated by Xiao and Bavel (2012) in their three studies on collective identity and identity threat. The authors found that threatening social situations were judged spatially closer than the non-threatening ones, reinforcing the idea of how distance perception serves the function of adjusting our behavior in relation to our social and physical environment. A final remark is on the application of the rules of physical and social distance not only to our egocentric perspective but also in the interpretation of social scenes in which more agents are interacting. The study of Zhou et al. (2019), among others, demonstrated that closer interpersonal distances, more direct interpersonal angles and more open postures, are all visual cues of ongoing interaction in a group of people. This study along with other experiments on how we interpret social scenes are described in greater detail in section 4.2 that is dedicated to the observation of multiple agents.

BOX 3 Definition of psychological distance.

This concept was first proposed by Trope and Liberman in their Construal Level Theory and was defined as the level of abstraction used to represent a phenomenon based on its temporal distance. Greater distance corresponds to greater abstraction. Now the theory includes other three categories of psychological distance: spatial, social, and hypothetical. As demonstrated also in this review, these four dimensions are strongly and systemically correlated with each other. Psychological distance is inevitably egocentric, the center is the self in the present, and it serves as a measure of the value attributed to the phenomenon of interest. Closer events/agents are perceived as more important and more likely to be acted on

The studies revised in this third section confirm that our body is the arena where we enact our own and other people's feelings, and the key to our complex social abilities. A summary of the critical studies in support of this concept is presented in Table 2. We can now finally explore in more detail how we use our vision to understand others, by describing those body indexes, such as posture and movement, that inform us on others' psychological state. This is the aim of the next section.

4. Reuniting visual perception and social cognition: The social body In neuroscientific research

Our interaction with others is substantially mediated by the observation of their behavior. As we just described above, we understand others' inner states by embodying their posture and expression, which elicit specific affective responses that we cognitively interpret and recognize (see subsection 3.2 for the embodied account of social cognition). In other words, it is by observing and mirroring the bodies of others that we gain insight of their inner states. The aim of this section is to provide the reader with an overview of the most recent techniques and to inspire new lines of research in visual social cognition. We report a summary of the techniques we describe in Table 3. We will distinguish between techniques that are used to examine posture, movement, and gait of individuals, from those that are used to inspect multiple agents interactions. We will end this methodological part by reporting some evidence demonstrating the social function of our vision, followed by a discussion on the importance of reintegrating the whole body in the study of emotion processing and social cognition.

4.1. Measures of posture, movement, and gait

The social cues we extract from other people's bodies are linked to their posture, movement and gait. Our emotions find expression not only by means of the facial muscles, but also in the way we position our limbs, shoulders and spine. For example, the curvature of the shoulders reflects behaviors of either closure or openness to the world, either avoidance or approaching attitude. Kinematics is another source of relevant information, and can be decomposed in different indexes: balance, movement and gait. Although the reliability of these body measures in predicting affective states is supported by an increasing number of studies, the tools and assessment methods to measure them are limited or underdeveloped in the empirical research. Here, we present a variety of instruments that can be used to quantify posture and movement.

4.1.1. Posture and gait

One of the most immediate and old ways of assessing gait speed and its characteristics is by videorecording people walking and analyzing photograms of the strides. A pioneer study was conducted in the 80s by Sloman et al. (1982), in which the authors assessed the gait in adults with depression using this method. The analysis of mobility in this clinical population showed that depression is associated with specific motor symptoms, such as slower movements and worse balance compared to healthy controls (Doumas et al., 2012; Belvederi Murri et al., 2020). The use of **electronic walkways** can provide a more accurate measure of the stride length and walking speed. Lemke et al. (2000) used a combination of photogrammetry and electronic walkway and confirmed the results found by Sloman proving a reduced stride length in depressed patients. Another study on depression (Hausdorff et al., 2004) adopted **pressure-sensitive shoe insoles** to check for variability in swing time, which resulted higher in the clinical population. To obtain indexes on the posture along with the walking characteristics, the use of **3D motion capture systems** can give more detailed information about head, e.g., position and movements, upper limbs swing, back curvature. For example, studies on depressed patients have shown a correlation between the severity of the depression and the thoracic curvature, supporting the idea that a slumped position can be associated with sadness and introversion (Belvederi Murri et al., 2020). The 3D motion capture system was applied by Angelini et al. (2020) on patients with multiple sclerosis. They identified diverse gait-based biomarkers using inertial motion sensors with the goal of improving the assessment of progressive multiple sclerosis (MS). The authors examined 15 gait measures and reported longer steps and stride duration, reduced regularity and higher instability in the walk of people with MS, when compared to healthy controls. The use of **wearable sensors** for the recording of the kinematics enables the collection of data outside the lab, the identification of a variety of gait parameters and the detection of biomarkers specific to different clinical conditions.

4.1.2. Balance

Balance can be assessed by means of force platforms (also known as stepping platforms) and in some cases the balance exercise performed during the execution of a working memory task can give information about how cognitive load can reduce balance skills (Doumas et al., 2012; Belvederi Murri et al., 2020).

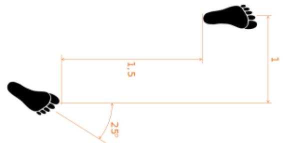
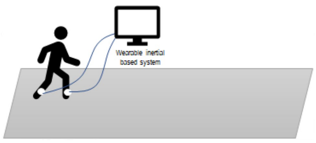
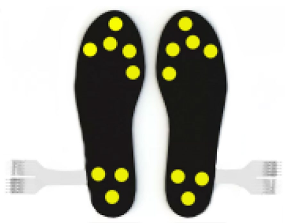
10.3389/fpsyg.2022.1028150

TABLE 2 Summary table of the studies cited in section 3.

Study	Sample	Stimuli	Design	Results and theoretical implications
Evidences for social embodiment				
Duclos et al. (1989)	$N = 74$ (43 females, undergraduate students)	Faces expressing sadness, anger, fear and disgust	Participants were induced to adopt expressions of fear, anger, disgust, and sadness by contracting or relaxing specific face muscles and then asked to rate their emotional state after having performed a conflicting task, designed to disguise the experiment's goal	Although participants were naïve to the aims of the experiment, they reported the highest rating of emotion in the condition in which they were expressing that emotion, giving strength to the idea that the activation of specific sensorimotor schemas elicits those embodied feelings
Rütgen et al. (2015)	$N = 102$ (70 females, mean age = 25)	Painful stimulation delivered to the participant or to another person	The researchers induced placebo analgesia induced in half of the participants and tested all participants' pain perception and empathy. fMRI activation and self-reported pain and empathy ratings were collected	Compared to the control group, participants with induced placebo analgesia showed reduced first-hand pain perception and reduced pain empathy, suggesting that pain empathy might be grounded in our own pain experiences
Lenzoni et al. (2020)	$N = 66$ (42 patients with myotonic dystrophy; 24 healthy controls)	FEAST-N, the subtest of the Facial Emotion Matching test, and BEAST-N, the subtest Body Emotion Matching	Emotional recognition ability was assessed in patients with myotonic dystrophy and compared with healthy controls	The clinical population performed significantly worse than the healthy controls, demonstrating a causal role of visuomotor abilities in emotion recognition
Ijzerman and Semin (2009)	$N = 33$ in Exp. 1; 52 in Exp. 2; 39 in Exp. 3	Inclusion of Other in Self (IOS) scale	Participants were asked to hold cold or hot beverages and rate the social proximity they felt with a known person of their choice using the IOS scale	Participants judged the distance between themselves and a known person as shorter when they were asked to hold a warm beverage, demonstrating an association between warmth and feelings of social proximity
Study	**Sample**	**Stimuli**	**Design**	**Results and theoretical implications**
Evidences for overlaps between social and physical space				
Bar-Anan et al. (2007)	$N = 10$ (6 females, undergraduate students)	Different words defining close or distant social affiliation depicted in different locations in a scenario	Participants had to indicate if the position of an item on the screen was proximal or distal, independently of the meaning of the word	Words were classified faster when psychological and spatial distances were matching, compared to when the two types of distances were incongruent
Kroczek et al. (2020)	$N = 36$ (18 females, mean age = 21.75)	Virtual agents represented in a VR scenario	Participants were instructed to approach the agents in the VR and start interacting as soon as the agent looked up at them. The interpersonal distance was varied by manipulating the distance at which agents reacted to the participant's approach. Arousal, valence, and realism rates were collected after each interaction, on a 1–100 points scale	Closer interpersonal distances were rated as more arousing, less pleasant, and less natural than longer distances
Yamakawa et al. (2009)	$N = 24$ (4 females, age range = 19–34 years)	Two inanimate objects whose relative physical positions could be inferred by texture and lighting cues (i.e., physical distance task). Pictures of two faces (i.e., social distance task)	In the physical distance task, participants had to judge which object was closer to them. In the social distance task, participants had to choose which person they felt more compatible with. fMRI data was acquired during both tasks	Results showed that the parietal cortex was activated in both tasks, suggesting a common neural substrate for the estimation of physical and social distances
Schiano Lomoriello et al. (2018)	$N = 34$ (23 females, mean age = 23)	Pictures of faces with neutral facial expression, receiving either a painful or a neutral stimulation. All faces were presented in the upright and inverted orientation and in two physical sizes, small and big, corresponding to a perceived far and close social distance	The perceived physical distance from the stimuli was manipulated through picture sizes. Participants were asked to assess the painfulness of the stimulation applied to each face presented. EEG data was collected during the task	ERPs modulations compatible with an empathic reaction were observed only for the group exposed to face stimuli appearing to be at a close social distance from the participants, i.e., big size pictures. This reaction was absent in the group exposed to smaller stimuli corresponding to face stimuli perceived from a far social distance
Zhou et al. (2019)	$N = 148$ in total across 7 experiments (mean age = 20)	Virtual avatars placed at different positions and with different face directions	Participants had to report if the avatars in the VR environment were interacting	Results showed that closer interpersonal distances, more direct interpersonal angles and more open postures, are all visual cues of ongoing interaction in a group of people

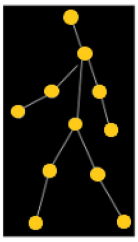
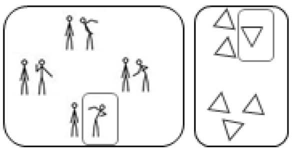
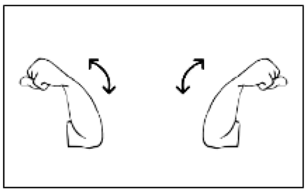
10.3389/fpsyg.2022.1028150

TABLE 3 Summary table of the techniques presented in sections "Measures of posture, movement, and gait", "Measures of observed social interactions", and " Measures of the observer".

	Measures of posture, movement, and gait"		
	Technique	Measures	Studies
 Image adapted from https://commons.wikimedia.org/wiki/File:Nachuo_sogi.svg	Stride and walk photogram analysis	Stride length and walking speed	Doumas et al. (2012); Sloman et al. (1982)
	Electronic walkway	Stride length and walking speed	Lemke et al. (2000)
	Pressure-sensitive shoe insoles	Stride length and walking speed	Hausdorff et al. (2004)
	3D motion capture system (inertial motion sensors)	Diverse gait-based biomarkers	Belvederi Murri et al. (2020)
	Wearable motion sensors	Gait parameters	Angelini et al. (2020)
 Image adapted from https://commons.wikimedia.org/wiki/File:AMTI_OPT464508_force_plate.png	Force platform	Balance	Belvederi Murri et al. (2020)

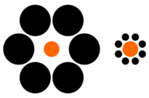
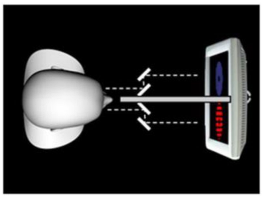
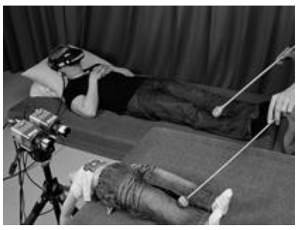
(Continued)

TABLE 3 (Continued)

	Point-light walker stimuli	Recognition of biological motion, also when it expresses emotional movements. Preference for detecting social agents facing us	Edey et al. (2017); Han et al. (2021); Miller and Saygin (2013)
	Measures of observed social interactions"		
	Technique	Measures	Studies
 Image adapted from https://commons.wikimedia.org/wiki/File:VRHeadset.png	Virtual reality	Observing or engaging in virtual social interaction. Manipulation of body schema	Kroczek et al. (2020); Yee et al. (2009); Zhou et al. (2019)
 Image adapted from https://pxhere.com/en/photo/1327220	Inversion effect	Configurational representation of the image in its typical display	Papeo and Abassi (2019); Stekelenburg and Gelder (2004)
	Object-inferiority effect	Faster detection of an object when represented outside a configurational figure that typically contains it	Papeo (2020)
	Videos of synchronous movement	Observation of social coordination as measure of social relationship	Macpherson et al. (2020)
	Measures of the observer		
	Technique	Measures	Studies
	Eye-tracker (eye movements and pupillometry)	Eye position, eye movements, pupil size. Saccades, fixation duration	Kret et al. (2013)

(Continued)

TABLE 3 (Continued)

 Image adapted from https://www.flickr.com/photos/sparkfun/8677911665 License CC-BY 4.0	Electromyography	Muscle activity, micro-movements	Kret et al. (2013)
 Ebbinghaus illusion. Ebbinghaus illusion Image adapted from https://www.google.com/search?q=ebbinghaus%20illusion&tbm=isch&tbs=il:cl&hl=en&sa=X&ved=0CAAQ1vwEahcKEwiYmtWH-M78AhUAAAAAHQAAAAAQAw&biw=1548&bih=937#imgrc=SCJW-hJd0NjdXM	Visual illusions	Magnitude of the illusion	Chouinard et al. (2018); King et al. (2017)
 Image adapted from https://commons.wikimedia.org/wiki/File:Binocular_rivalry_Experiment.png License CC-BY 4.0	Binocular rivalry	Perceptual dominance of one of two stimuli competing to reach visual awareness	Anderson et al. (2011)
 Image adapted from https://picryl.com/media/goggles-8b0197	Skewed goggles	Effects of magnification of objects	Linkenauger et al. (2010)
 Image adapted with permission from van der Hoort et al. (2011) License CC-BY 4.0	Body-swap illusion	Influence of our body size on the perception of our physical and social environment	van der Hoort et al. (2011)

This dual task approach can be implemented also during other movement assessments, as it is effective in detecting how cognitive load influences motor skills in clinical populations.

4.1.3. Observing moving bodies: Point-light walker stimuli

Obtaining gait and balance measurements can be exploited also to study our social vision: Edey et al. (2017) investigated the walking speed of participants and tested whether participants used their own kinematics as a reference to judge the affective states of point-light walker (PLW) stimuli. These visual stimuli are animations composed solely of points that have been previously attached to the joints of a moving person and extrapolated by the video recording of the scene. Following the idea that our own kinematics influences our perception of emotional movements in others, the authors manipulated the speed and posture of the artificial walker in order to elicit anger, happiness or sadness. As expected, there was a modulatory effect of the participant's movements on the emotion recognition: people judged less intensely emotions similar to their own walking pace. In other words, participants who walked with greater speed rated high-velocity emotions (e.g., anger) as less intense relative to low-velocity emotions (e.g., sadness). This finding is in agreement with the theories of embodied social cognition reviewed above (see subsection 3.2). Point-light stimuli in motion has been used also to investigate the detection of biological motion in relation to measures of social cognition. Using this approach, it has been shown that higher scores in social cognition tests were linked to high accuracy in biological motion detection (Miller and Saygin, 2013). Moreover, PLW stimuli have been used to measure our promptness in seeing social agents facing us compared to facing away based on the perceived social relevance. The level of social relevance was manipulated by changing distance, speed and size of the PLW, based on the assumption that people perceived as nearer, faster and bigger have more social relevance than those perceived as farther, slower and smaller. Therefore, the likelihood of initiating an interaction with them increases. PLW stimuli are particularly suited to measure differences in seeing people facing toward us or away thanks to their ambiguous in-depth orientation (Han et al., 2021). Findings from these studies show once again that social factors have a clear impact on visual processes, further supporting our hypothesis of a deep link between low-level feature detection and high-level social cognition.

4.2. Measures of observed social interactions

Beside the interface with a single person, our social life is mainly constituted by crowded situations in which multiple agents engage complex interactions with each other. A relatively new branch of study is focusing on the perception of the relations between social entities and proposes that our interpretation of social events draws upon a configurational recognition process. Recent findings suggest a specific sensitivity of the visual system for the spatial relationship between multiple social agents, such as interpersonal distance and angle of a facing dyad (Papeo, 2020). One of the requirements for a successful social interaction is, indeed, a face-to-face position between the agents. This implies that seeing two people facing each other makes us assume an ongoing interaction (Zhou et al., 2019). Based on the data from their experiments with virtual reality (VR), Zhou et al. (2019) created a computational model of the social interaction field, which they describe as the area surrounding each of us within which we can start interacting with other people. Similarly, to a gravitational field, the social interaction field can inform us on the strength of the social interaction between two people based on their physical distance and positions in space.

4.2.1. Facing dyads

We believe, akin to other authors (Papeo, 2020), that our visual system is tuned for the recognition of social interactions around us, allowing us a fast detection of social groups. The aggregation of multiple elements (individuals) in a unitary piece of visual information (group) can facilitate and fasten the representation of the crowded scene we are observing (Zhou et al., 2019). By means of the inversion effect described above, Papeo and Abassi (2019) demonstrated that facing dyads are processed as unitary perceptual objects. They found that the inversion effect was greater for the facing dyads compared to non-facing ones. To recall the definition of this phenomenon, a greater inversion effect implies greater visual sensitivity, in this case supporting the hypothesis of a visual attunement for interacting agents. It should be noted that the two bodies inversion effect is not observed for non-human or for human-object dyads. Another signature of this visual grouping is the object-inferiority effect that has been observed in the visual search through a crowd: the dyad as a whole is detected faster than the single objects within it, but only when the agents are facing each other (Papeo et al., 2019).

4.2.2. Synchronicity as a measure of social relation

We extract social information about an ongoing interaction also from the observation of the agents' movements. In this case, it is the level of interpersonal coordination that informs us about the cooperative or hostile tones of a social interaction. A higher synchronization of the dyad's movements signals coalition rather than opposition (Macpherson et al., 2020).

Miles et al. (2009) used stick figures and sounds of footstep to simulate two people walking together with different gait patterns and found that when rhythms of walking synchrony were out of phase, these were associated with a lower level of relationship. A similar study showed that social factors, such as the skin-color, can influence the perception of **synchronous movements** (Macpherson et al., 2020).

4.3. Measures of the observer

After having presented numerous techniques suitable for measuring body position and movement that can be observed in one or multiple social agents, we want to offer an overview of methodologies that can be used to analyze the observer's behavior.

For a comprehensive approach in visual social cognition, it is important to appropriately combine measures that quantify the observed social cues with measures describing the observer's state, at a cognitive, visual and physiological level.

4.3.1. Eye-trackers and other physiological indexes

Since we are concentrating on the visual aspects of social cognition, studying the eye and the way we visually scan the social scene is almost imperative. Although the most immediate way to study gaze behavior is by means of **eye-trackers**, the range of methodologies does not limit to this one. Other important indicators of social cognition - ideally to be combined with the gaze measurements - are those linked to the automatic mimicry involved in the process of emotion recognition. In this case, the focus is on the muscles and posture of the observer and the mirroring reflexes can be recorded through the application of **sensible electrodes on expression muscles**.

Kret et al. (2013) collected eye movements, pupil size, and facial muscles activity, while participants were performing an emotion discrimination task with full body and face stimuli. They used full body images from the BEAST (Bodily Expressive Action Stimulus Test, de Gelder and Van den Stock, 2011) database. The database includes stimuli representing emotions in whole-body figures, and in this experiment they were presented in association with congruent or incongruent facial expressions. Eye movements were recorded with a wearable eye-tracking device, and their analysis revealed that participants looked longer at faces than at bodies, and the same applied for happy versus angry/fearful postures, with longer fixation duration for negative body expressions compared to positive ones. Furthermore, negative emotion expressions correlated with activity in the observers' corrugator, and happy expressions with zygomaticus' activation. The corrugator showed more responsiveness for bodies compared to faces, whilst a reversed pattern was observed for the zygomaticus. These findings nicely dovetail with theories of embodied social cognition, supporting the preconscious activation of the expressive muscles matching the emotion observed. Other implicit indexes of emotional response are detectable by physiological measures, such as skin conductance and heart rate variability (for reviews on emotion measures see: Mauss and Robinson, 2009; Egger et al., 2019).

4.3.2. Measures of visual perception

Visual illusions can provide useful insights on the interplay between high and low order processes in the perception of an image. We will describe the application of this kind of tool in more detail in the section dedicated to the clinical studies (subsection 5). Assessment of visual awareness can be provided with dichoptic stimulation (i.e., simultaneous presentation of different stimuli to the two eyes), which provokes the phenomenon of binocular rivalry, i.e., the alternation in the perception of two different images presented to each eye.

In an experiment by Anderson et al. (2011), a paradigm with binocular rivalry was used to examine the influence of the affective state of a perceiver on the visual awareness of the stimulus presented. The potential of this technique lies in the fact that the two visual inputs presented to the two visual hemispheres compete for perceptual dominance; the selective criteria are driven by top-down processes and this allows us to determine how our internal state influences visual awareness. The authors first manipulated the participants' affective states by showing them emotional images, and subsequently asked them to perform the binocular rivalry tasks. These consisted of a neutral stimulus (e.g., a house) presented in competition with a socially relevant stimulus (facial expressions) and participants had to report what they were seeing and for how long. Results confirmed the hypothesis of the authors, for which the affective state of the viewer biases the contents of visual awareness. In fact, the social stimuli were always dominant in the image perception, and this effect was maximized when participants were asked to watch a set of stimuli inducing unpleasant emotions. These findings show how binocular rivalry can be used to explore the process of sensory selection behind our conscious experience of the world and support the role of top-down modulation on our visual perception.

4.3.3. Body schema manipulations and their effects on personality

We would like to dedicate a short section also to some methodologies used to investigate the influence of our body size and appearance on our perception of the physical and social environments. In an attempt to tackle this issue, previous research has relied upon illusions, which were generated either by magnifying or minimizing the objects in the visual field, or by inducing the sensation of having a shrinked or gigantified body. In a study, Linkenauger et al. (2010) observed the rescaling effects induced by placing one's own hand close to objects, whose size was distorted by means of skewed goggles. The presence of a personal body part canceled out the magnifying or minimizing effect of the illusion. In another study, van der Hoort et al. (2011) generated a deeper manipulation of the body schema, referred to as body-swap illusion, by touching a part of the participants' body and showing them a video of the same tactile stimulation being performed on a mannequin of different dimensions. Although the retinal images remained identical, perceiving a different size of the body changed the estimates of size and distance of objects present in the scene (van der Hoort et al., 2011).

Lastly, virtual avatars can alter our self-representation. Yee et al. (2009) had people interact in virtual environments with avatars of different dimensions, and found that taller and attractive avatars outperformed shorter avatars in the online game "World of Warcraft." The authors attributed the better performance to an increased self-esteem and confidence linked to the height and attractiveness of the characters, showing that an avatar's appearance can influence a user's behavior in an online environment. The effect was transposed also outside the virtual environment: in a second experiment, a VR session in which participants had either a tall or short avatar was followed by a face-to-face interaction during which a negotiation task was performed. It turned out that people that had embodied a tall avatar were more likely to act unfairly to gain more profit and less prone to accept transactions against their interests than participants in short avatars. These studies reveal the potential of VR as a promising technique not only for their observational scope but also as promising intervention tools in clinical settings.

4.4. Our eyes at the service of emotion recognition and social communication

The inextricable connection between vision and social cognition has biological plausibility: the human eye (Schutt et al., 2015). Our eyes seem to have evolved to serve the fine and complex phenomenon of human communication and the development of social skills through our gaze-following abilities. These abilities are favored by certain characteristics specific to our species: the white sclera of our eyes and the high contrast in eye and facial skin coloration. Indeed, only in humans the outline of the eyes and the position of the iris are so clearly visible, conveying information on where the others are looking (Proffitt, 2006; Tomasello et al., 2007). In a study by Tomasello et al. (2007), gaze-following behavior was studied in both primates and infants. This behavior is based on cues coming from head orientation or from eyes direction. In this study, an experimenter sat in front of the ape or the child to be tested and looked up at the ceiling in different modalities: only with the eyes while keeping the head in a frontal position, bending backward the neck and facing the ceiling with the eye closed, or with face and eyes both looking up. Results showed a preference in infants for eye direction cues independently from the head orientation, while great apes relied mostly on head direction cues, suggesting that humans are more attuned to the eyes than our closest primate relatives (the great apes) are (Tomasello et al., 2007). The unique features of the human eye probably represent the key to mechanisms of shared attention which are at the basis of the human propensity for cooperation and coordination (Shepherd, 2010).

The dominance of the visual system in human communication is evident also in the automatic and fast identification of faces and bodies even in the most complex scenarios. As anticipated in subsection 3.3, we are able to detect conspecifics and evaluate the spatial relations between them in a very rapid and preconscious way. The preconscious nature of emotion processing guided by our vision has been investigated in different studies in which patients with lesions to the primary visual cortex could still perform task of emotion discrimination, without visual awareness of the stimulus (Morris et al., 2001; Pegna et al., 2005; Tamietto and de Gelder, 2008; de Gelder, 2009). For instance, Pegna et al. (2005) collected data from one patient who became cortically blind as a consequence of two strokes that destroyed his visual cortices bilaterally. Different visual discrimination tasks showed a small capacity to discriminate emotional social stimuli (expressive faces), whilst no sensitivity was observed for different kinds of stimuli (e.g., neutral faces, animals). Similar results were obtained by Morris et al. (2001) on a patient with right hemianopia due to a left occipital lobe damage. Taken together, these findings suggest a strong connection between visual inputs and subcortical structures, aimed at providing an automatic discrimination of salient, emotional stimuli.

In summary, the human eye does not serve solely vision, but social communication and emotion processing as well (Proffitt, 2006). Once again, the distinction between cognitive and sensory processes becomes even more blurred, reinforcing those models that postulate early influences of our sociality on our perceptual systems.

4.5. Expressive bodies, not just faces: Reintegrating the whole body in the study of visual social cognition

Just a decade ago, only less than 5% of the experimental production had considered the inclusion of the whole body as stimuli in their design (de Gelder, 2009). By now, the situation has seen little changes: Witkower et al. (2021) argued also for the need of further investigation on how bodies convey social information. In their study on a culturally-isolated population of Nicaragua, they have shown effective recognition of bodily basic expressions of sadness, anger, and fear, in the members of this society, providing evidence for the universality of these bodily displays. Indeed, a growing body of evidence is testifying that body expressions are recognized automatically and effectively, in the same specialized manner that characterizes the innate predisposition to face perception (for a review see, Quinn and Macrae, 2011).

Behavioral and physiological data have proven that emotion recognition relies considerably on the observation of body expressions. For instance, Kret et al. (2013) presented participants with *ad-hoc* images of body emotional postures associated with congruent or incongruent facial expressions (fear or happiness). Results revealed that the recognition of the emotion expressed by the face was influenced by the emotion expressed by the body. Response time increased with incongruent stimuli, while it decreased when face and body were expressing the same feeling. Stekelenburg and Gelder (2004) examined the electrophysiological correlates of the inversion effect, a well-known phenomenon in facial perception for which people take longer to recognize faces presented upside-down compared to any other object presented in the same fashion. The effect is explained by assuming a configural representation for the identification of faces, which fastens their detection when they appear in the expected upright position but slows it down when inverted. By using EEG, the authors showed that the same ERP component, namely the N170, was evoked both by faces and bodies presented upside down, but not by pictures of inverted objects (e.g., shoes), suggesting a configural coding of bodies' images similar to the one underlying face perception. Studies that investigated functional connectivity between brain areas active during recognition of bodily expressions (Peelen and Downing, 2005; van de Riet et al., 2009) showed the activation of the same areas that typically respond to face stimuli. These areas appeared to be only a part of the broader network involved in body stimuli processing, which also includes the supratemporal sulcus, the middle temporal/middle occipital gyrus, the superior occipital gyrus and the parieto-occipital sulcus.

These dedicated mechanisms behind body perception support the importance of recognizing bodily expressions in our everyday life.

5. Insights from clinical studies

Finally, lesion and clinical studies are valuable in the examination of causal relationships between social and perceptual processes. Neuromuscular diseases, for instance, can provide insights into the relation between emotion recognition in others and impaired sensorimotor skills. Such an example is the study of Lenzoni et al. (2020) on myotonic dystrophy described in section 3.2. Along with motor impairments, clinical categories in which social deficits represent the major symptomatology can be studied for investigating the connection between social cognition and perceptual anomalies. Autism spectrum disorder and schizophrenia offer the unique opportunity to examine possible links between deficits in social abilities and altered visual perception - that are typically observed in these disorders (Butler et al., 2008; King et al., 2017; Robertson and Baron-Cohen, 2017; Chung and Son, 2020).

In a review by King et al. (2017), perceptual abnormalities in schizophrenia have been revised through the analysis of studies on visual illusions and their effects on this clinical population. Perceptual illusions are widely used in vision studies, in that they allow to disentangle the mechanisms underlying visual processing. For example, the Ebbinghaus illusion (for which a target item looks smaller or bigger by effect of contextual cues) can be modulated by the effects of prior knowledge and culture on visual perception, which means that it is a distortion linked to top-down processing of the visual inputs. From the literature reviewed in King et al. (2017), it emerged that it is this kind of high-level integration that seems to be systematically altered in people with schizophrenia. In fact, they tend to show a reduced susceptibility for high-level illusions, suggesting that abnormalities in visual perception might depend on deficits in the cognitive/perceptual communication at the basis of perceptual awareness. As the same authors suggest, it would be helpful to study these processes not only in isolation but also applying converging techniques to investigate the reciprocal links between higher and lower processes with ecologically valid designs. In this way, it might be possible to explore more in depth the connections between inferential top-down aspects of visual perception and the ability of recognizing social cues from observing other people.

In a similar fashion, different perceptual styles in autism were examined in a study by Chouinard et al. (2018), in which sensory integration was investigated again by means of visual illusions. In this case, the Shepard illusion was tested in autistic and typically developing individuals while their eye movements were recorded. In contrast with the authors' expectations, no difference was found in saccades and scene exploration between the two groups, although the clinical population experienced a weaker illusion than the healthy controls. These results can be explained by differences in high-level visual integration, instead of anomalies in earlier stages of perception (e.g., spatial exploration, saccade velocity and frequency). As for schizophrenia, the empirical data suggest that top-down inferences might be reduced in people with autism, bringing to higher objectivity in perceiving the world as it really is, which in turn leads to a diminished sensitivity to visual illusions. Once more, further research is needed in order to establish the relationship between these perceptual anomalies and deficits in higher order social cognition.

6. Conclusion and future directions of research

Our review strives to encourage the application of multimodal integrative approaches in cognitive, social and affective neuroscience and to inspire further research aimed at discovering the intertwined connection between social cognition and visual perception.

We highlighted the tight relationship between visual perception and social cognition. Specifically, we aimed at unveiling the role of the body as the starting point for the construction of our perception, also when it comes to social perception. Firstly, we compared research modalities that can be adopted in the exploration of these constructs, and remarked on the necessity of moving from isolationist unimodal approaches to integrative multimodal perspectives. Subsequently, we presented abundant evidence for the rooting of social cognition in bodily expressions, as defined by theories of embodied social cognition. Finally, we described a common mechanism at the basis of specific aspects of social cognition and spatial behavior: an overlapping neural network for the perception of both physical and social distances. It appears that we recruit networks dedicated to the processing of physical distances to map our social environment, strengthening the dependence of higher order social abilities on lower representational systems closer to perceptual networks.

This visual-social interface is at play also in processes of emotion recognition, as we rely on visual cues collected through the scrutiny of the other's body. Again, the body acts as a middle ground where our vision and our social abilities can encounter.

We also described some of the body features we observe to assess the other's state, and reviewed instruments and techniques useful in quantifying these indexes for research applications. We described different ways to measure posture, balance, and gait, as meaningful indicators of emotional states. Also the interaction among multiple agents was covered in the methodological section, by providing examples of studies that examined interpersonal distance and synchronization as a hint for understanding the quality of the relationship. Finally, we described different ways to analyze the observer's body, such as the detection of micro-movements underlying a first stage of emotional contagion, or the visual exploration of the stimuli by means of eye-tracking devices. The body is the key for any level of reciprocal understanding, and combining the study of bodily expressions with the analysis of visual behavior can be beneficial for the development of a detailed model of social cognition.

6.1. Limitations

A major limitation of this review is that we were unable to extensively cover the whole literature on the topics discussed here. Due to space constraints, sometimes we failed to establish a balance between sources supporting and those opposing a particular view. In this paragraph, we would like to at least introduce the reader to the ongoing debate on what we believe to be one of the core themes of the review: the embodied account of cognition.

Probably the main criticism against embodiment theories is that they disregard any mental constructs of the perceived events. In an interesting paper by Borg (2018), the main argument against these theories relies on the absence of mental state attribution in action understanding, as we predict or explain the behavior of others by adopting what she refers to as a 'smart behavior reading'.

According to this model, action understanding depends directly on non-mentalistic interactive embodied practices (e.g., sensitivity to physical context and bodily motions) rather than on our ability to understand and interact with others. As such, the smart behavior reading account does not take into consideration the individuality of the observed person and all the information we might have about their personality, or life circumstances, which would enable us to predict completely different outcomes of their actions. Hence, the slow, controlled and demanding characteristics required by a mentalistic interpretation of other people's behavior are 'deflated' by the embodied accounts of social cognition in favor of a fast, effortless and automatic behavior reading. Nonetheless, we also believe that mental state attribution cannot be completely disregarded by models of action understanding without generating gaps/errors in interpretation, and only an integrative approach which combines both smart behavior tracking and mental state attribution would enable successful action prediction.

Another critique of the embodied accounts comes from studies by Mahon and Caramazza (2008) and Caramazza et al. (2014). Specifically, these authors argue against the explanation of previous electrophysiological research only in light of an embodied view of action understanding processes. The authors claim that the empirical evidence provided by neuroimaging research can be equally used in support of disembodied views of conceptual representations or at least they do not necessarily discard them. Motor and sensory activation during action representation can be seen as part of a cascade process that propagates through qualitatively different levels of processing. Nevertheless, Caramazza et al., also acknowledge the authenticity of sensory-motor activation during action observation or evocation, and propose instead a middle-ground theory that combines together the abstract and symbolic levels of some concepts with the more embodied instantiation of online conceptual processing (Mahon and Caramazza, 2008; Caramazza et al., 2014).

6.2. Future directions

This review is not the first nor the only one that points out the importance of the body in shaping cognitive processes (see for example, Harris et al., 2015). Nonetheless, it strives to have both theoretical and practical implications. The summary of different methodologies and instruments measuring body indexes and psychological responses can be a useful source of information for researchers interested in conducting studies using the paradigms described here.

For instance, physiological measures may be included in studies of visual perception to explore whether the awareness of the body to external visual stimuli precedes their conscious appraisal or vice versa, deepening our understanding of implicit and explicit processing of emotional stimuli. Moreover, recentering clinical investigations on the body indices could be particularly relevant in those syndromes characterized both by deficits in social cognition and visual perception, such as autism or schizophrenia. Critically, in these syndromes verbal communication can be severely impaired or hardly accessible. Future studies may investigate the use of the body for patterns of interactions with the external world (e.g., postures, gait) as a way to access these clinical conditions. For example, the numerous studies on synchronization as an indirect measure of relational quality could inspirate group or couple exercises aimed at eliciting cooperative behavior.

A clearer understanding of the interplay between visual perception and social cognition might also help the development of novel clinical treatments or cognitive training that takes into account perceptual alterations in order to improve social abilities. For instance, differences in visual processing in individuals with autism (e.g., less semantic-oriented, more detail-oriented) have been identified in the literature and linked to altered abilities in social cognition, such as emotion recognition. A novel training approach might aim at reinforcing global visual processing and this in turn might lead to an improvement in emotion recognition in this population.

Again, systematic investigations of physiological measures linked to the observation of others' bodies may represent a turning point in the study of empathy and reciprocal understanding of others' inner states. These investigations may represent an innovative way to assess empathy implicitly whilst overcoming issues associated with self-reported measures both in health participants and patients.

Finally, we hope that this review will further strengthen multi-level and multi-approach explanations of cognitive processes that attempt to promote the integration of embodiment theories with more traditional cognitivist approaches. We believe that these trends are already present in the literature, and that they will lead to a comprehensive framework and an integrated perspective on perception and cognition in the years to come.

In summary, we highlighted how the visual perception of our social and physical environments is mediated by bodily processes and expressions. The body is behind any possible interaction with our surroundings and our conspecifics. Observing other people's bodies informs us on their psychological state and allows emotional sharing and understanding. Our body determines the way we look at the world surrounding us.

Ethics statement

Written informed consent was obtained from the individual(s) for the publication of any potentially identifiable images or data included in this article.

Author contributions

CD, IS, and FM contributed equally to the literature search, literature discussion, and writing of the manuscript. All authors contributed to the article and approved the submitted version.

Acknowledgments

The authors thank the Department of Psychology and Cognitive Science (DiPSCo) of the University of Trento for the financial support provided for the submission of this paper.

Conflict of interest

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

Publisher's note

References

Abraham, A., Werning, M., Rakoczy, H., von Cramon, D. Y., and Schubotz, R. I. (2008). Minds, persons, and space: an fMRI investigation into the relational complexity of higher-order intentionality. *Conscious. Cogn.* 17, 438–450. doi: 10.1016/j.concog.2008.03.011

Ainley, V., Apps, M. A. J., Fotopoulou, A., and Tsakiris, M. (2016). Bodily precision: a predictive coding account of individual differences in interoceptive accuracy. *Philos. Trans. R. Soc. B Biol. Sci.* 371:20160003. doi: 10.1098/rstb.2016.0003

Anderson, E., Siegel, E. H., and Barrett, L. F. (2011). What you feel influences what you see: the role of affective feelings in resolving binocular rivalry. *J. Exp. Soc. Psychol.* 47, 856–860. doi: 10.1016/j.jesp.2011.02.009

Angelini, L., Hodgkinson, W., Smith, C., dd, J. M., Sharrack, B., Mazzà, C., et al. (2020). Wearable sensors can reliably quantify gait alterations associated with disability in people with progressive multiple sclerosis in a clinical setting. *J. Neurol.* 267, 2897–2909. doi: 10.1007/s00415-020-09928-8

Balcetis, E., and Dunning, D. (2006). See what you want to see: motivational influences on visual perception. *J. Pers. Soc. Psychol.* 91, 612–625. doi: 10.1037/0022-3514.91.4.612

Balcetis, E., and Dunning, D. (2009). Wishful seeing: more desired objects are seen as closer. *Psychol. Sci.* 21, 147–152. doi: 10.1177/0956797609356283

Bar-Anan, Y., Liberman, N., Trope, Y., and Algom, D. (2007). Automatic processing of psychological distance: evidence from a Stroop task. *J. Exp. Psychol. Gen.* 136, 610–622. doi: 10.1037/0096-3445.136.4.610

Barrett, L. F., and Satpute, A. B. (2019). Historical pitfalls and new directions in the neuroscience of emotion. *Neurosci. Lett.* 693, 9–18. doi: 10.1016/j.neulet.2017.07.045

Bastiaansen, J. A. C. J., Thioux, M., and Keysers, C. (2009). Evidence for mirror systems in emotions. *Philos. Trans. R. Soc. Lond. Ser. B Biol. Sci.* 364, 2391–2404. doi: 10.1098/rstb.2009.0058

Beckes, L., and Coan, J. A. (2011). Social Baseline Theory: The Role of Social Proximity in Emotion and Economy of Action. *Soc. Personal. Psychol. Compass* 5, 976–988.

Belvederi Murri, M., Triolo, F., Coni, A., Tacconi, C., Nerozzi, E., Escelsior, A., et al. (2020). Instrumental assessment of balance and gait in depression: a systematic review. *Psychiatry Res.* 284:112687. doi: 10.1016/j.psychres.2019.112687

Berti, A., and Frassinetti, F. (2000). When far becomes near: remapping of space by tool use. *J. Cogn. Neurosci.* 12, 415–420. doi: 10.1162/089892900562237

Bhalla, M., and Proffitt, D. R. (1999). Visual–motor recalibration in geographical slant perception. *J. Exp. Psychol. Hum. Percept. Perform.* 25, 1076–1096. doi: 10.1037/0096-1523.25.4.1076

Blanke, O., Slater, M., and Serino, A. (2015). Behavioral, neural, and computational principles of bodily self-consciousness. *Neuron* 88, 145–166. doi: 10.1016/j.neuron.2015.09.029

Bonini, L., Rotunno, C., Arcuri, E., and Gallese, V. (2022). Mirror neurons 30 years later: implications and applications. *Trends Cogn. Sci.* 26, 767–781. doi: 10.1016/j.tics.2022.06.003

Borg, E. (2018). On deflationary accounts of human action understanding. *Rev. Philos. Psychol.* 9, 503–522. doi: 10.1007/s13164-018-0386-3

Brockmole, J. R., Davoli, C. C., Abrams, R. A., and Witt, J. K. (2013). The World Within Reach: Effects of Hand Posture and Tool Use on Visual Cognition. *Curr. Dir. Psychol. Sci.* 22, 38–44 doi: 10.1177/0963721412465065.

Bruner, J. S., and Goodman, C. C. (1947). Value and need as organizing factors in perception. *J. Abnorm. Soc. Psychol.* 42, 33–44. doi: 10.1037/h0058484

Butler, P. D., Silverstein, S. M., and Dakin, S. C. (2008). Visual perception and its impairment in schizophrenia. *Biol. Psychiatry* 64, 40–47. doi: 10.1016/j.biopsych.2008.03.023

Caramazza, A., Anzellotti, S., Strnad, L., and Lingnau, A. (2014). Embodied cognition and Mirror neurons: a critical assessment. *Annu. Rev. Neurosci.* 37, 1–15. doi: 10.1146/annurev-neuro-071013-013950

Changizi, M. A., and Hall, W. G. (2001). Thirst modulates a perception. *Perception* 30, 1489–1497. doi: 10.1068/p3266

Chouinard, P. A., Royals, K. A., Landry, O., and Sperandio, I. (2018). The Shepard illusion is reduced in children with an autism Spectrum disorder because of perceptual rather than attentional mechanisms. *Front. Psychol.* 9:2452. doi: 10.3389/fpsyg.2018.02452

Chung, S., and Son, J.-W. (2020). Visual perception in autism Spectrum disorder: a review of neuroimaging studies. *J. Korean Acad. Child Adolesc. Psychiatry* 31, 105–120. doi: 10.5765/jkacap.200018

Cole, S., Balcetis, E., and Dunning, D. (2013). Affective Signals of Threat Increase Perceived Proximity. *Psychol. Sci.* 24, 34–40. doi: 10.1177/0956797612446953

de Gelder, B. (2009). Why bodies? Twelve reasons for including bodily expressions in affective neuroscience. *Philos. Trans. R. Soc. B Biol. Sci.* 364, 3475–3484. doi: 10.1098/rstb.2009.0190

De Gelder, B., and Van den Stock, J. (2011). The bodily expressive action stimulus test (BEAST). Construction and validation of a stimulus basis for measuring perception of whole body expression of emotions. *Front. Psychol.* 2:181. doi: 10.3389/fpsyg.2011.00181

de Gelder, B., Van den Stock, J., Meeren, H. K. M., Sinke, C. B. A., Kret, M. E., and Tamietto, M. (2010). Standing up for the body: Recent progress in uncovering the networks involved in the perception of bodies and bodily expressions. *Neurosci. Biobehav. Rev.* 34, 513–527. doi: 10.1016/j.neubiorev.2009.10.008

de Pinedo García, M. (2020). Ecological psychology and Enactivism: a normative way out from ontological dilemmas. *Front. Psychol.* 11:1637. doi: 10.3389/fpsyg.2020.01637

Doerrfeld, A., Sebanz, N., and Shi, M. (2012). Expecting to lift a box together makes the load look lighter. *Psychol. Res.* 76, 467–475. doi: 10.1007/s00426-011-0398-4

Doumas, M., Smolders, C., Brunfaut, E., Bouckaert, F., and Krampe, R. T. (2012). Dual task performance of working memory and postural control in major depressive disorder. *Neuropsychology* 26, 110–118. doi: 10.1037/a0026181

Duclos, S. E., Laird, J. D., Schneider, E., Sexter, M., Stern, L., and Van Lighten, O. (1989). Emotion-specific effects of facial expressions and postures on emotional experience. *J. Pers. Soc. Psychol.* 57, 100–108. doi: 10.1037/0022-3514.57.1.100

Edey, R., Yon, D., Cook, J., Dumontheil, I., and Press, C. (2017). Our own action kinematics predict the perceived affective states of others. *J. Exp. Psychol. Hum. Percept. Perform.* 43, 1263–1268. doi: 10.1037/xhp0000423

Egger, M., Ley, M., and Hanke, S. (2019). Emotion recognition from physiological signal analysis: a review. *Electron. Notes Theor. Comput. Sci.* 343, 35–55. doi: 10.1016/j.entcs.2019.04.009

Fabbri-Destro, M., and Rizzolatti, G. (2008). Mirror neurons and Mirror Systems in Monkeys and Humans. *Physiology* 23, 171–179. doi: 10.1152/physiol.00004.2008

Fairhurst, M., Fairhurst, K., Berna, C., and Tracey, I. (2012). An fMRI study exploring the overlap and differences between neural representations of physical and recalled pain. *PLoS One* 7:e48711. doi: 10.1371/journal.pone.0048711

Firestone, C., and Scholl, B. J. (2016). Cognition does not affect perception: evaluating the evidence for "top-down" effects. *Behav. Brain Sci.* 39:e229. doi: 10.1017/S0140525X15000965

Fodor, J. A. (1983). *The modularity of mind.* Cambridge, Mass, London: MIT Press.

Forkmann, K., Wiech, K., Sommer, T., and Bingel, U. (2015). Reinstatement of pain-related brain activation during the recognition of neutral images previously paired with nociceptive stimuli. *Pain* 156, 1501–1510. doi: 10.1097/j.pain.0000000000000194

Friston, K. (2018). Does predictive coding have a future? *Nat. Neurosci.* 21, 1019–1021. doi: 10.1038/s41593-018-0200-7

Friston, K., and Kiebel, S. (2009). Predictive coding under the free-energy principle. *Philos. Trans. R. Soc. B Biol. Sci.* 364, 1211–1221. doi: 10.1098/rstb.2008.0300

Fuchs, T. (2020). The circularity of the embodied mind. *Front. Psychol.* 11:707. doi: 10.3389/fpsyg.2020.01707

Gallese, V. (2007). Before and below 'theory of mind': embodied simulation and the neural correlates of social cognition. *Philos. Trans. R. Soc. B Biol. Sci.* 362, 659–669. doi: 10.1098/rstb.2006.2002

Gallese, V. (2009). Mirror neurons, embodied simulation, and the neural basis of social identification. *Psychoanal. Dialogues* 19, 519–536. doi: 10.1080/10481880903231910

Gibson, J. J. (2014). *The Ecological Approach to Visual Perception: Classic Edition 1st Edn.* New York: Psychology Press.

Goldstone, R. L., and Barsalou, L. W. (1998). Reuniting perception and conception. *Cognition* 65, 231–262. doi: 10.1016/S0010-0277(97)00047-4

Gross, E. B., and Proffitt, D. (2013). The economy of social resources and its influence on spatial perceptions. *Front. Hum. Neurosci.* 7:772. doi: 10.3389/fnhum.2013.00772

Hall, E. T. (1963). A system for the notation of Proxemic behavior. *Am. Anthropol.* 65, 1003–1026. Available at: http://www.jstor.org/stable/668580. doi: 10.1525/aa.1963.65.5.02a00020

Hall, E. T., Birdwhistell, R. L., Bock, B., Bohannan, P., Diebold, A. R., Durbin, M., et al. (1968). Proxemics [and comments and replies]. *Curr. Anthropol.* 9, 83–108. Available at: http://www.jstor.org/stable/2740724. doi: 10.1086/200975

Hamilton, A. F. de C., Kessler, K., and Creem-Regehr, S. H. (2014). Perspective taking: building a neurocognitive framework for integrating the "social" and the "spatial." *Front. Hum. Neurosci.* 8:35. doi: 10.3389/fnhum.2014.00403

Han, Q., Wang, Y., Jiang, Y., and Bao, M. (2021). The relevance to social interaction modulates bistable biological-motion perception. *Cognition* 209:104584. doi: 10.1016/j.cognition.2021.104584

Harber, K. D., Yeung, D., and Iacovelli, A. (2011). Psychosocial resources, threat, and the perception of distance and height: support for the resources and perception model. *Emotion* 11:1080.

Harris, L. R., Carnevale, M. J., D'Amour, S., Fraser, L. E., Harrar, V., Hoover, A. E. N., et al. (2015). How our body influences our perception of the world. *Front. Psychol.* 6:819. doi: 10.3389/fpsyg.2015.00819

Hausdorff, J. M., Peng, C.-K., Goldberger, A. L., and Stoll, A. L. (2004). Gait unsteadiness and fall risk in two affective disorders: a preliminary study. *BMC Psychiatry* 4:39. doi: 10.1186/1471-244X-4-39

Heft, H. (2020). Ecological psychology and Enaction theory: divergent groundings. *Front. Psychol.* 11:991. doi: 10.3389/fpsyg.2020.00991

Henderson, M. D., Fujita, K., Trope, Y., and Liberman, N. (2006). Transcending the "here": the effect of spatial distance on social judgment. *J. Pers. Soc. Psychol.* 91, 845–856. doi: 10.1037/0022-3514.91.5.845

Ijzerman, H., and Semin, G. R. (2009). The thermometer of social relations: mapping social proximity on temperature. *Psychol. Sci.* 20, 1214–1220. doi: 10.1111/j.1467-9280.2009.02434.x

Keysers, C., Kaas, J. H., and Gazzola, V. (2010). Somatosensation in social perception. *Nat. Rev. Neurosci.* 11, 417–428. doi: 10.1038/nrn2833

Keysers, C., Paracampo, R., and Gazzola, V. (2018). What neuromodulation and lesion studies tell us about the function of the mirror neuron system and embodied cognition. *Curr. Opin. Psychol.* 24, 35–40. doi: 10.1016/j.copsyc.2018.04.001

King, D. J., Hodgekins, J., Chouinard, P. A., Chouinard, V.-A., and Sperandio, I. (2017). A review of abnormalities in the perception of visual illusions in schizophrenia. *Psychon. Bull. Rev.* 24, 734–751. doi: 10.3758/s13423-016-1168-5

Kret, M. E., Stekelenburg, J. J., Roelofs, K., and de Gelder, B. (2013). Perception of face and body expressions using electromyography pupillometry and gaze measures. *Front. Psychol.* 4:28. doi: 10.3389/fpsyg.2013.00028

Kroczek, L. O. H., Pfaller, M., Lange, B., Müller, M., and Mühlberger, A. (2020). Interpersonal distance during real-time social interaction: insights from subjective experience, behavior, and physiology. *Front. Psych.* 11:561. doi: 10.3389/fpsyt.2020.00561

Lamm, C., Rütgen, M., and Wagner, I. C. (2019). Imaging empathy and prosocial emotions. *Neurosci. Lett.* 693, 49–53. doi: 10.1016/j.neulet.2017.06.054

Landau, M. J., Meier, B. P., and Keefer, L. A. (2010). A metaphor-enriched social cognition. *Psychol. Bull.* 136, 1045–1067. doi: 10.1037/a0020970

Lemke, M. R., Wendorff, T., Mieth, B., Buhl, K., and Linnemann, M. (2000). Spatiotemporal gait patterns during over ground locomotion in major depression compared with healthy controls. *J. Psychiatr. Res.* 34, 277–283. doi: 10.1016/S0022-3956(00)00017-0

Lenzoni, S., Bozzoni, V., Burgio, F., de Gelder, B., Wennberg, A., Botta, A., et al. (2020). Recognition of emotions conveyed by facial expression and body postures in myotonic dystrophy (DM). *Cortex* 127, 58–66. doi: 10.1016/j.cortex.2020.02.005

Lindblom, J. (2020). A radical reassessment of the body in social cognition. *Front. Psychol.* 11:987. doi: 10.3389/fpsyg.2020.00987

Linkenauger, S., and Proffitt, D. (2008). The effect of intention and bodily capabilities on the perception of size. *J. Vis.* 8:620. doi: 10.1167/8.6.620

Linkenauger, S. A., Ramenzoni, V., and Proffitt, D. R. (2010). Illusory Shrinkage and Growth: Body-Based Rescaling Affects the Perception of Size. *Psychol. Sci. 21*, 1318–1325 doi: 10.1177/0956797610380700

Macpherson, M. C., Fay, N., and Miles, L. K. (2020). Seeing synchrony: a replication of the effects of task-irrelevant social information on perceptions of interpersonal coordination. *Acta Psychol.* 209:103140. doi: 10.1016/j.actpsy.2020.103140

Maggio, M. G., Piazzitta, D., Andaloro, A., Latella, D., Sciarrone, F., Casella, C., et al. (2022). Embodied cognition in neurodegenerative disorders: what do we know so far? A narrative review focusing on the mirror neuron system and clinical applications. *J. Clin. Neurosci.* 98, 66–72. doi: 10.1016/j.jocn.2022.01.028

Mahon, B. Z., and Caramazza, A. (2008). A critical look at the embodied cognition hypothesis and a new proposal for grounding conceptual content. *J. Physiol. Paris* 102, 59–70. doi: 10.1016/j.jphysparis.2008.03.004

Mauss, I. B., and Robinson, M. D. (2009). Measures of emotion: a review. *Cogn. Emot.* 23, 209–237. doi: 10.1080/02699930802204677

Meconi, F., Doro, M., Schiano Lomoriello, A., Mastrella, G., and Sessa, P. (2018). Neural measures of the role of affective prosody in empathy for pain. *Sci. Rep.* 8:291. doi: 10.1038/s41598-017-18552-y

Meconi, F., Linde-Domingo, J. S., Ferreira, C., Michelmann, S., Staresina, B., Apperly, I. A., et al. (2021). EEG and fMRI evidence for autobiographical memory reactivation in empathy. *Hum. Brain Mapp.* 42, 4448–4464. doi: 10.1002/hbm.25557

Miles, L. K., Nind, L. K., and Macrae, C. N. (2009). The rhythm of rapport: interpersonal synchrony and social perception. *J. Exp. Soc. Psychol.* 45, 585–589. doi: 10.1016/j.jesp.2009.02.002

Miller, L. E., and Saygin, A. P. (2013). Individual differences in the perception of biological motion: links to social cognition and motor imagery. *Cognition* 128, 140–148. doi: 10.1016/j.cognition.2013.03.013

Morris, J. S., DeGelder, B., Weiskrantz, L., and Dolan, R. J. (2001). Differential extrageniculostriate and amygdala responses to presentation of emotional faces in a cortically blind field. *Brain* 124, 1241–1252. doi: 10.1093/brain/124.6.1241

Mukamel, R., Ekstrom, A. D., Kaplan, J., Iacoboni, M., and Fried, I. (2010). Single-neuron responses in humans during execution and observation of actions. *Curr. Biol.* 20, 750–756. doi: 10.1016/j.cub.2010.02.045

Nakayama, K. (2011). "Introduction: Vision Going Social." in *The science of social vision. Vol. 7.* eds. R. B. Adams, R. B. Adams Jr, N. Ambady, K. Nakayama and S. Shimojo (Oxford university press).

Niedenthal, P. M., Barsalou, L. W., Winkielman, P., Krauth-Gruber, S., and Ric, F. (2005). Embodiment in attitudes, social perception, and emotion. *Personal. Soc. Psychol. Rev.* 9, 184–211. doi: 10.1207/s15327957pspr0903_1

Niedenthal, P., and Wood, A. (2019). Does emotion influence visual perception? Depends on how you look at it. *Cogn. Emot.* 33, 77–84. doi: 10.1080/02699931.2018.1561424

Noel, J.-P., Blanke, O., and Serino, A. (2018). From multisensory integration in peripersonal space to bodily self-consciousness: from statistical regularities to statistical inference. *Ann. N. Y. Acad. Sci.* 1426, 146–165. doi: 10.1111/nyas.13867

Oishi, S., Schiller, J., and Gross, E. B. (2013). Felt understanding and misunderstanding affect the perception of pain, slant, and distance. *Soc. Psychol. Personal. Sci.* 4, 259–266. doi: 10.1177/1948550612453469

Otten, M., Seth, A. K., and Pinto, Y. (2017). A social Bayesian brain: how social knowledge can shape visual perception. *Brain Cogn.* 112, 69–77. doi: 10.1016/j.bandc.2016.05.002

Papeo, L. (2020). Twos in human visual perception. *Cortex* 132, 473–478. doi: 10.1016/j.cortex.2020.06.005

Papeo, L., and Abassi, E. (2019). Seeing social events: the visual specialization for dyadic human–human interactions. *J. Exp. Psychol. Hum. Percept. Perform.* 45, 877–888. doi: 10.1037/xhp0000646

Papeo, L., Goupil, N., and Soto-Faraco, S. (2019). Visual search for people among people. *Psychol. Sci.* 30, 1483–1496. doi: 10.1177/0956797619867295

Paradiso, E., Gazzola, V., and Keysers, C. (2021). Neural mechanisms necessary for empathy-related phenomena across species. *Curr. Opin. Neurobiol.* 68, 107–115. doi: 10.1016/j.conb.2021.02.005

Parkinson, C., and Wheatley, T. (2013). Old cortex, new contexts: re-purposing spatial perception for social cognition. *Front. Hum. Neurosci.* 7:645. doi: 10.3389/fnhum.2013.00645

Peelen, M. V., and Downing, P. E. (2005). Selectivity for the human body in the fusiform gyrus. *J. Neurophysiol.* 93, 603–608. doi: 10.1152/jn.00513.2004

Pegna, A. J., Khateb, A., Lazeyras, F., and Seghier, M. L. (2005). Discriminating emotional faces without primary visual cortices involves the right amygdala. *Nat. Neurosci.* 8, 24–25. doi: 10.1038/nn1364

Popova, Y. B., and Rączaszek-Leonardi, J. (2020). Enactivism and ecological psychology: the role of bodily experience in agency. *Front. Psychol.* 11:841. doi: 10.3389/fpsyg.2020.539841

Proffitt, D. R. (2006). Embodied perception and the economy of action. *Perspect. Psychol. Sci.* 1, 110–122. doi: 10.1111/j.1745-6916.2006.00008.x

Proffitt, D., and Baer, D. (2020). *Perception: How Our Bodies Shape Our Minds* Minds. St. Martin's Press.

Proffitt, D., and Linkenauger, S. (2013). "Perception viewed as a phenotypic expression" in *Action science: Foundations of an emerging discipline*. eds. W. Prinz, M. Beisert, and A. Herwig (MIT Press), 171–197.

Quinn, K. A., and Macrae, C. N. (2011). The face and person perception: insights from social cognition. *Br. J. Psychol.* 102, 849–867. doi: 10.1111/j.2044-8295.2011.02030.x

Rao, R. P. N., and Ballard, D. H. (1999). Predictive coding in the visual cortex: a functional interpretation of some extra-classical receptive-field effects. *Nat. Neurosci.* 2, 79–87. doi: 10.1038/4580

Read, C., and Szokolszky, A. (2020). Ecological psychology and Enactivism: perceptually-guided action vs. sensation-based Enaction 1. *Front. Psychol.* 11:270. doi: 10.3389/fpsyg.2020.01270

Robertson, C. E., and Baron-Cohen, S. (2017). Sensory perception in autism. *Nat. Rev. Neurosci.* 18, 671–684. doi: 10.1038/nrn.2017.112

Rütgen, M., Seidel, E.-M., Pletti, C., Riečanský, I., Gartus, A., Eisenegger, C., et al. (2018). Psychopharmacological modulation of event-related potentials suggests that first-hand pain and empathy for pain rely on similar opioidergic processes. *Neuropsychologia* 116, 5–14. doi: 10.1016/j.neuropsychologia.2017.04.023

Rütgen, M., Seidel, E.-M., Silani, G., Riečanský, I., Hummer, A., Windischberger, C., et al. (2015). Placebo analgesia and its opioidergic regulation suggest that empathy for pain is grounded in self pain. *Proc. Natl. Acad. Sci.* 112, E5638–E5646. doi: 10.1073/pnas.1511269112

Sato, A., Matsuo, A., and Kitazaki, M. (2019). Social contingency modulates the perceived distance between self and other. *Cognition* 192:104006. doi: 10.1016/j.cognition.2019.06.018

Schiano Lomoriello, A., Meconi, F., Rinaldi, I., and Sessa, P. (2018). Out of sight out of mind: perceived physical distance between the observer and someone in pain shapes Observer's neural empathic reactions. *Front. Psychol.* 9:1824. doi: 10.3389/fpsyg.2018.01824

Schnall, S., Harber, K. D., Stefanucci, J. K., and Proffitt, D. R. (2008). Social support and the perception of geographical slant. *J. Exp. Soc. Psychol.* 44, 1246–1255. doi: 10.1016/j.jesp.2008.04.011

Schutt, R. K., Seidman, L. J., and Keshavan, M. S. (Eds.). (2015). *Social neuroscience: Brain. mind. and society.* Harvard University Press.

Serino, A. (2019). Peripersonal space (PPS) as a multisensory interface between the individual and the environment, defining the space of the self. *Neurosci. Biobehav. Rev.* 99, 138–159. doi: 10.1016/j.neubiorev.2019.01.016

Serino, A., Alsmith, A., Costantini, M., Mandrigin, A., Tajadura-Jimenez, A., and Lopez, C. (2013). Bodily ownership and self-location: components of bodily self-consciousness. *Conscious. Cogn.* 22, 1239–1252. doi: 10.1016/j.concog.2013.08.013

Sessa, P., Meconi, F., and Han, S. (2014). Double dissociation of neural responses supporting perceptual and cognitive components of social cognition: evidence from processing of others' pain. *Sci. Rep.* 4:424. doi: 10.1038/srep07424

Seth, A. K. (2013). Interoceptive inference, emotion, and the embodied self. *Trends Cogn. Sci.* 17, 565–573. doi: 10.1016/j.tics.2013.09.007

Seth, A., Suzuki, K., and Critchley, H. (2012). An interoceptive predictive coding model of conscious presence. *Front. Psychol.* 2:395. doi: 10.3389/fpsyg.2011.00395

Shepherd, S. (2010). Following gaze: gaze-following behavior as a window into social cognition. *Front. Integr. Neurosci.* 4:5. doi: 10.3389/fnint.2010.00005

Sloman, L., Berridge, M., Homatidis, S., Hunter, D., and Duck, T. (1982). Gait patterns of depressed patients and normal subjects. *Am. J. Psychiatry* 139, 94–97. doi: 10.1176/ajp.139.1.94

Stekelenburg, J., and Gelder, B. (2004). The neural correlates of perceiving human bodies: an ERP study on the body-inversion effect. *Neuroreport* 15, 777–780. doi: 10.1097/00001756-200404090-00007

Sugovic, M., Turk, P., and Witt, J. K. (2016). Perceived distance and obesity: It's what you weigh, not what you think. *Acta Psychol.* 165, 1–8. doi: 10.1016/j.actpsy.2016.01.012

Sun, C., Chen, J., Chen, Y., and Tang, R. (2021). The influence of induced emotions on distance and size perception and on the grip scaling during grasping. *Front. Psychol.* 12:651885. doi: 10.3389/fpsyg.2021.651885

Takahashi, K., Meilinger, T., Watanabe, K., and Bülthoff, H. H. (2013). Psychological influences on distance estimation in a virtual reality environment. *Front. Hum. Neurosci.* 7:580. doi: 10.3389/fnhum.2013.00580

Tamietto, M., and de Gelder, B. (2008). Affective blindsight in the intact brain: neural interhemispheric summation for unseen fearful expressions. *Neuropsychologia* 46, 820–828. doi: 10.1016/j.neuropsychologia.2007.11.002

Teneggi, C., Canzoneri, E., di Pellegrino, G., and Serino, A. (2013). Social modulation of peripersonal space boundaries. *Curr. Biol.* 23, 406–411. doi: 10.1016/j.cub.2013.01.043

Tomasello, M., Hare, B., Lehmann, H., and Call, J. (2007). Reliance on head versus eyes in the gaze following of great apes and human infants: the cooperative eye hypothesis. *J. Hum. Evol.* 52, 314–320. doi: 10.1016/j.jhevol.2006.10.001

Trope, Y., and Liberman, N. (2010). Construal-level theory of psychological distance. *Psychol. Rev.* 117, 440–463. doi: 10.1037/a0018963

Valenti, J. J., and Firestone, C. (2019). Finding the "odd one out": memory color effects and the logic of appearance. *Cognition* 191:103934. doi: 10.1016/j.cognition.2019.04.003

van de Riet, W. A. C., Grezes, J., and de Gelder, B. (2009). Specific and common brain regions involved in the perception of faces and bodies and the representation of their emotional expressions. *Soc. Neurosci.* 4, 101–120. doi: 10.1080/17470910701865367 van der Hoort, B., Guterstam, A., and Ehrsson, H. H. (2011). Being Barbie: the size of one's own body determines the perceived size of the world. *PloS One* 6:e20195. doi: 10.1371/journal.pone.0020195

Varela, F. J., Thompson, E., and Rosch, E. (1991). *The embodied mind: Cognitive science and human experience*. Cambridge, MA, US: The MIT Press.

Witkower, Z., Hill, A. K., Koster, J., and Tracy, J. L. (2021). Beyond face value: evidence for the universality of bodily expressions of emotion. *Affect. Sci.* 2, 221–229. doi: 10.1007/s42761-021-00052-y

Xiao, Y. J., and Bavel, J. J. V. (2012). See your friends close and your enemies closer. *Personal. Soc. Psychol. Bull.* 38, 959–972. doi: 10.1177/0146167212442228

Yamakawa, Y., Kanai, R., Matsumura, M., and Naito, E. (2009). Social distance evaluation in human parietal cortex. *PLoS One* 4:e4360. doi: 10.1371/journal.pone.0004360

Yamazaki, Y., Hashimoto, T., and Iriki, A. (2009). The posterior parietal cortex and non-spatial cognition. *Biol. Rep.* 1:74. doi: 10.3410/B1-74

Yee, N., Bailenson, J. N., and Ducheneaut, N. (2009). The proteus effect: implications of transformed digital self-representation on online and offline behavior. *Commun. Res. 36*, 285–312. doi: 10.1177/0093650208330254

Zadra, J. R., Weltman, A. L., and Proffitt, D. R. (2016). Walkable distances are bioenergetically scaled. *J. Exp. Psychol. Hum. Percept. Perform.* 42, 39–51. doi: 10.1037/xhp0000107

Zaki, J., and Ochsner, K. N. (2012). The neuroscience of empathy: progress, pitfalls and promise. *Nat. Neurosci.* 15, 675–680. doi: 10.1038/nn.3085

Zhou, C., Han, M., Liang, Q., Hu, Y.-F., and Kuai, S.-G. (2019). A social interaction field model accurately identifies static and dynamic social groupings. *Nat. Hum. Behav.* 3, 847–855. doi: 10.1038/s41562-019-0618-2