(DIS)EMBODIED PERCEPTION OF THE SELF AND OTHER - INTERDISCIPLINARY PERSPECTIVES FROM SCIENCE & ARTS

EDITED BY: Anna Ciaunica, Vittorio Gallese, Hanne De Jaegher, Aikaterini (Katerina) Fotopoulou, Frank Röhricht, David Harris Smith and Tereza Stehlikova

PUBLISHED IN: Frontiers in Psychology

frontiers Research Topics

frontiers

ISSN 1664-8714
ISBN 978-2-88974-205-9
DOI 10.3389/978-2-88974-205-9

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

(DIS)EMBODIED PERCEPTION OF THE SELF AND OTHER - INTERDISCIPLINARY PERSPECTIVES FROM SCIENCE & ARTS

Topic Editors:
Anna Ciaunica, University of Porto, Portugal
Vittorio Gallese, University of Parma, Italy
Hanne De Jaegher, University of the Basque Country, Spain
Aikaterini (Katerina) Fotopoulou, University College London, United Kingdom
Frank Röhricht, University of Essex, United Kingdom
David Harris Smith, McMaster University, Canada
Tereza Stehlikova, University of Westminster, United Kingdom

Citation: Ciaunica, A., Gallese, V., De Jaegher, H., Fotopoulou, A. K., Röhricht, F., Smith, D. H., Stehlikova, T., eds. (2022). (Dis)Embodied Perception of the Self and Other - Interdisciplinary Perspectives From Science & Arts.
Lausanne: Frontiers Media SA. doi: 10.3389/978-2-88974-205-9

Table of Contents

04 ***The Active Room: Freud's Office and the Egyptian Tomb***
Julia K. Schroeder

15 ***How Does Embodying a Transgender Narrative Influence Social Bias? An Explorative Study in an Artistic Context***
Marte Roel Lesur, Sonia Lyn and Bigna Lenggenhager

23 ***Losing Ourselves: Active Inference, Depersonalization, and Meditation***
George Deane, Mark Miller and Sam Wilkinson

38 ***Pain Asymbolia as Depersonalization for Pain Experience. An Interoceptive Active Inference Account***
Philip Gerrans

48 ***"Dance Like Nobody's Watching": Exploring the Role of Dance-Based Interventions in Perceived Well-Being and Bodily Awareness in People With Parkinson's***
Rebecca Hadley, Olivia Eastwood-Gray, Meryl Kiddier, Dawn Rose and Sonia Ponzo

59 ***Into Your (S)Kin: Toward a Comprehensive Conception of Empathy***
Tue Emil Öhler Søvsø and Kirstin Burckhardt

74 ***Why Build a Robot With Artificial Consciousness? How to Begin? A Cross-Disciplinary Dialogue on the Design and Implementation of a Synthetic Model of Consciousness***
David Harris Smith and Guido Schillaci

88 ***Whatever Next and Close to My Self—The Transparent Senses and the "Second Skin": Implications for the Case of Depersonalization***
Anna Ciaunica, Andreas Roepstorff, Aikaterini Katerina Fotopoulou and Bruna Petreca

102 ***Bodies in the Novel* Infinite Jest**
Ana Chapman, Silvia Chapman and Stephanie Cosentino

111 ***[Re]moving Bodies – A Shared Diminished Reality Installation for Exploring Relational Movement***
Julien Laroche, Loup Vuarnesson, Alexandra Endaltseva, Joseph Dumit and Asaf Bachrach

frontiers in Psychology

HYPOTHESIS AND THEORY
published: 28 July 2020
doi: 10.3389/fpsyg.2020.01547

The Active Room: Freud's Office and the Egyptian Tomb

*Julia K. Schroeder**

Department of German, University of California, Berkeley, Berkeley, CA, United States

The present study examines the striking similarities between the architectural design and spatial composition of the ancient Egyptian tomb and Sigmund Freud's office at Berggasse 19 in Vienna, Austria. I argue that the Egyptian tomb elements represented within Freud's office permitted the enclosed space to play an active role in his psychoanalysis sessions. I supplement this argument by analyzing the office's spatial and architectural arrangements in relation to ancient Egyptian architectural frameworks, psychoanalytic container theory (Freud, Danze, and Quinodoz), and Freud's archeological metaphor model. This study contributes to the greater body of work on architecture as an active entity, psychoanalysis, and ancient Egyptian history.

Keywords: Sigmund Freud, office design, psychoanalysis, work environment, active room, therapy, extended mind, ancient Egypt

OPEN ACCESS

Edited by:
Hanne De Jaegher,
University of the Basque Country, Spain

Reviewed by:
Martin Lee Gledhill,
University of Bath, United Kingdom
Peggy Deamer,
Yale University, United States

***Correspondence:**
Julia K. Schroeder
juliaschroeder@berkeley.edu

Specialty section:
This article was submitted to Theoretical and Philosophical Psychology, a section of the journal Frontiers in Psychology

Received: *05 March 2020*
Accepted: *09 June 2020*
Published: *28 July 2020*

Citation:
Schroeder JK (2020) The Active Room: Freud's Office and the Egyptian Tomb. Front. Psychol. 11:1547. doi: 10.3389/fpsyg.2020.01547

> *I then made some short observations upon the psychological differences between the conscious and unconscious, and upon the fact that everything conscious was subject to a process of wearing-away, while what was unconscious was relatively unchangeable; and I illustrated my remarks by pointing to the antiques standing in my room. They were, in fact, I said, only objects found in a tomb, and their burial had been their preservation.*
>
> – Sigmund Freud, *Notes Upon a Case of Obsessional Neurosis*

FREUD'S OFFICE AND THE EGYPTIAN TOMB

In his 1937 essay "Constructions in Analysis," Sigmund Freud related King Tutankhamun's well-preserved Egyptian tomb to his psychoanalytic practice. Upon closer inspection, it appears that Freud's office floor plan nearly matches Egyptian 18th Dynasty tomb floor plans, including that of King Tutankhamun. In the tomb and office, the space that held the body was composed of conjoining, externally sealed off rooms. Representations of essential Egyptian tomb elements and furniture were also included inside Freud's enclosed, tomb-like office.

The office's spatial and physical design appears to have permitted the patient to participate in free association and, in turn, overcome past traumas. The Egyptian tomb's spatial and physical design was considered to permit the mummy's spirit to transcend the body and transition into the afterlife. Following contemporary, psychoanalytic "active room" theory (Danze, 2005) and "active container" theory (Quinodoz, 1992), Freud's office and the ancient Egyptian tomb were active rooms–closed-off spaces that significantly contributed to the contained individual's psychic transformation.

The office's Egyptian elements and tomb-like architecture appear to illuminate the profound influence Egyptian culture and civilization had on Freud's psychoanalytic setting and practice. Freud's iconic, tomb-like office laid the foundation for the contemporary emphasis on psychoanalytic office design and architecture. The analyst's room retains its active role in psychoanalysis sessions. The present analysis contributes to the greater body of work on architecture as an active entity, psychoanalysis, and ancient Egyptian history. Furthermore, this

study builds upon the article "Berggasse 19: Inside Freud's Office" in which Diana Fuss and Joel Sanders posit Freud's "labor of self-entombment" within his office space (Fuss and Sanders, 2015, p. 3). The present article extends beyond Freud's self-entombment and posits that the tomb-like elements and architecture in Freud's office, moreover, played a central role in facilitating the transformation and liberation of the patient's psyche.

ARCHEOLOGY AND ARCHITECTURE

Freud's office floor plan bears a strong resemblance to 18th Dynasty Egyptian tomb floor plans, including that of King Tutankhamun (see **Figure 1**). Freud's office consisted of his consulting room and study, which formed a single room. In King Tutankhamun's tomb, the burial chamber and treasury formed a single room. In Freud's office and the Egyptian tomb, an open doorway connected the two spaces. Each room was a contained space complete with an open portal at its center (see **Figure 2**).

A comparison of the tomb and office floor plans may be viewed as Freud's consulting room being in the position of the tomb's burial chamber. In the burial chamber, the mummy lay in the sarcophagus. In Freud's consulting room, the patient lay on the couch. Freud's study was in the position of the tomb treasury. The mummy's personal belongings and furniture were kept in the treasury. The majority of Freud's Egyptian treasures and furniture were kept in his study. Commenting on Freud's Egyptian treasures, Fuss and Sanders write that Freud assembled "antiquities that would transform his office into a veritable tomb" (Fuss and Sanders, 2015, p. 3). Augmenting this claim, in his memoirs, Freud's patient Sergei Pankejeff remarked that the office "in no way reminded one of a doctor's office . . . Here were all kinds of statuettes and other unusual objects, which even a layman recognized as archeological findings from ancient Egypt . . . Everything here . . . contributed to one's feeling . . . of being sheltered" (Gardiner, 1973, p. 139).

The floor plans of Freud's office and the Egyptian tomb reveal that each space's external doors remained sealed. Freud's office and the Egyptian tomb appear to be closed-off containers that significantly contributed to psychic and spiritual transformations, respectively. Each space was defined by its external, closed-off form and internal transformative function. These characteristics resonate with the contemporary psychoanalytic active container framework.

ACTIVE CONTAINER

The Egyptian tomb and Freud's office represent active containers. According to Swiss psychoanalyst Danielle Quinodoz's (1992) article "The Psychoanalytic Setting as the Instrument of the Container Function," the psychoanalyst's office is an "active container." In other words, the room "is not an inert vessel . . . but an active container which interacts dynamically with the [psychoanalysis] process" (Quinodoz, 1992, p. 629). The active container's external features enable the patient to "gain access to the new relational world in which the unconscious psychical mechanisms begin to come alive and in which internal psychical reality becomes as real as external reality" (Quinodoz, 1992, p. 629). Freud's office and the ancient Egyptian tomb were enclosed spaces that opened up new relational worlds for the mummy and patient. Each room held the body and permitted the mind to undergo a liberating, psychic transformation.

In the tomb, the mummy's body rested in the sarcophagus as its spirit journeyed into the afterlife. In the office, the patient rested on Freud's couch and journeyed into the realm of the unconscious. In the present study, the Egyptian afterlife serves as a metaphor for the unconscious processes accessed by the patient during free association. Each room liberated the mind, while the housed body remained physically in a resting position. The tomb-like architecture and spatial composition of Freud's office stimulated the resting patient's psyche and physical body. Freud himself drew connections between ancient tombs and psychoanalysis in his writings on the archeological metaphor.

FREUD'S ARCHEOLOGICAL METAPHOR

Freud's comparisons between psychoanalysis and archeology formed the basis of his archeological metaphor. This metaphor further reveals how his contained, tomb-like office played an active role in liberating the patient's mind. In his writings, Freud linked his psychic excavations to tomb archeologist Heinrich Schliemann's physical excavations. Drawing upon Schliemann's excavation of Mycenae, Freud wrote in his essay "Female Sexuality": "our insight into this early, pre-Oedipus, phase in girls comes to us as a surprise, like the discovery, in another field, of the Minoan–Mycenaean civilizations behind the civilization of Greece" (Freud, 1931, p. 226). Freud claimed the archeologist and the psychoanalyst were united by their shared mission to discover and preserve the past. The tomb and office actively aided the archeologist and the psychoanalyst in this shared endeavor. Freud's passion for archeology resurfaces in his writings, office architecture, and book collection.

ARCHEOLOGIST HEINRICH SCHLIEMANN'S INFLUENCE ON FREUD

Freud cherished his copy of Heinrich Schliemann's Mycenae and Tiryns archeology reports, which "were inscribed by him with the date of their purchase in the 1890s" (Freud Museum London, 2018, p. 1). In Schliemann's (1878) publication entitled *Mycenae: A Narrative of Researches and Discoveries at Mycenae and Tiryns*, he described his excavation of these two ancient Greek cities. Throughout his descriptions of his 1878 Greek tomb excavations in Mycenae, Schliemann draws comparisons between Greek and Egyptian tomb architecture and elements. As stated in the preface written by the prime minister of England, William Ewart Gladstone, Schliemann's work "affords a new indication of prehistoric relations between Mycenae and Egypt" (Schliemann, 1878, p. 7). In the 10th chapter on "the five Royal Tombs at Mycenae," Schliemann notes the "whole household

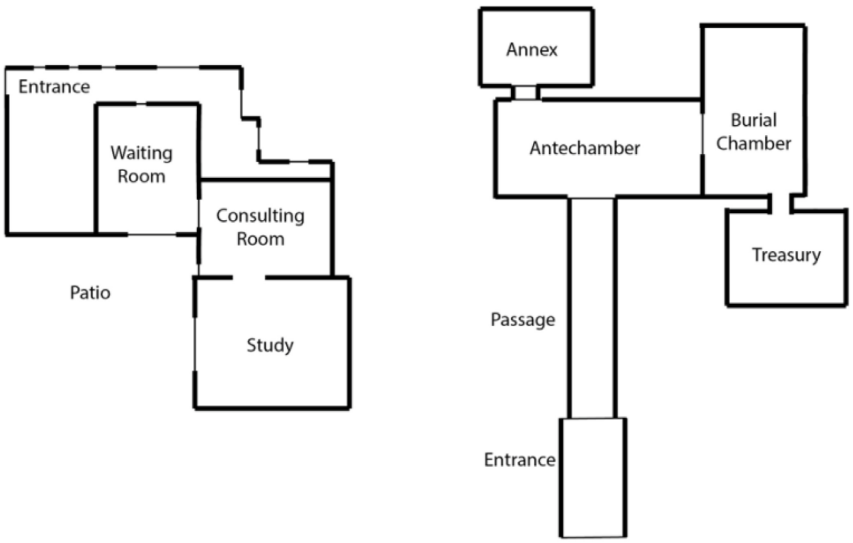

FIGURE 1 | The **left** image shows a floor plan of Sigmund Freud's office at Berggasse 19 in Vienna, Austria. Note that the consulting room and study are conjoined by an open doorway. The door between the waiting room and the consulting room is sealed. The **right** image shows a floor plan of King Tutankhamun's tomb. Note that the burial chamber and treasury are conjoined by an open doorway. The door between the antechamber and burial chamber is sealed. (**Figure 1** by Julia K. Schroeder).

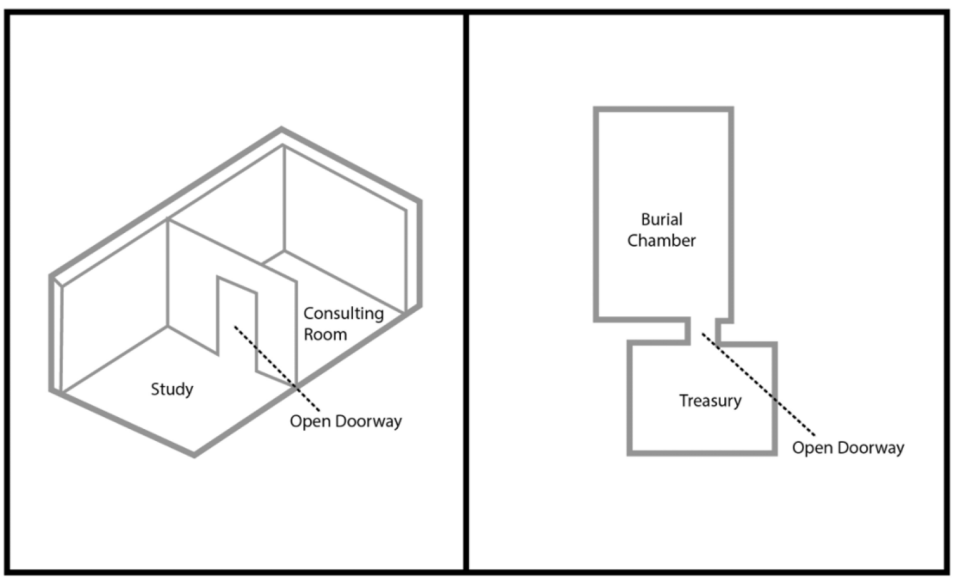

FIGURE 2 | The **left** image shows a three-dimensional model of Freud's office at Berggasse 19. The **right** image shows a model of King Tutankhamun's tomb in Egypt's Valley of the Kings. Note that each space is a single room. (**Figure 2** by Julia K. Schroeder).

furniture, copper kettles, drinking vessels, and so forth of a rich warrior" were found in the tombs (Schliemann, 1878, p. 349). Commenting on these findings, Schliemann writes: "I hardly think it necessary to remind the reader of the custom of ancient Egypt of burying the dead with treasures, for all the collections of Egyptian antiquities in the world are produced from Egyptian tombs" (Schliemann, 1878, p. 349). In the Egyptian tomb and Freud's office, the antiquities appeared to play an essential part in activating the rooms.

FREUD'S TOMB-LIKE OFFICE

In 1909, Freud wrote that the antiquities found in his study and conjoint consulting room were "only objects found in a tomb" (Freud, 1909, p. 176). Freud might well have actively created a tomb-like atmosphere within his office. By 1938, Freud had collected more than 2,000 ancient objects. Approximately half of these objects were from ancient Egypt. In the 2005 article "An Architect's View of Introspective Space: The Analytic Vessel,"

architect and professor Elizabeth A. Danze delineates how such objects contribute to the experience within the psychoanalyst's active container, which she terms an "active room." Danze claims that the active room's "objects, items, and art . . . might provide opportunities for association . . . These items are not neutral, but rather hold and present opportunities . . . they are ripe for symbolic interpretation" (Danze, 2005, p. 123). Freud's comparisons between psychoanalysis and tomb archeology reveal the symbolic value of his antiquities and their active role in his practice.

In "Constructions in Analysis," Freud related the resurfacing of intact memories in psychoanalysis to the excavation of King Tutankhamun's intact tomb. Freud claimed analysts "are regularly met by a situation which with the archaeological object occurs only in such rare circumstances as those of Pompeii or of the tomb of Tutankhamun. All of the essentials are preserved; even things that seem completely forgotten are present" (Freud, 1937, p. 260). The tomb-like office not only resembled King Tutankhamun's intact tomb but also provided optimal conditions for the analyst and patient to access and recover the buried past. The psyche's buried past may be drawn to the surface within the active room.

SPACE AS EXTENSION OF THE PSYCHIC APPARATUS

In each active room, space appears to represent an extension of the psyche. In 1938, Freud wrote that "space may be the projection of the extension of the psychical apparatus. No other derivation is possible . . . Psyche is extended" (Freud, 1938, p. 23). Freud seemed to conceptualize space as an active entity. The active room's space may be conceived as forming a discursive force field around the analyst and the analysand. The force field was charged by the psychic energy of past and present, conscious and unconscious thought processes. Here, it is important to note that in Freud's writings, the unconscious is often "ignorant of time, conserving its objects like an Egyptian tomb" (Deleuze and Guattari, 2004, p. 106). Much like the sealed off Egyptian tomb, the office's "psychic force field" appears to have been reinforced by physical walls and closed doors.

In his *Introductory Lectures on Psychoanalysis*, Freud writes: "I have had the ordinary door between my waiting room and my office doubled and strengthened by a covering of felt . . . it is in his [the patient's] interest not to be listened to while he is talking to the physician" (Freud, 1917, pp. 196–197). In Freud's office, closing the soundproof doors activated the room by providing the patient with the security he needed to draw his unconscious memories to the surface. This security was enhanced by Freud in 1907, when he moved his study and consulting room to the back of the Berggasse 19 apartment. This further isolated the office from the Freud family's living quarters. Contained by the office's walls and closed, soundproof doors, the patient lying on Freud's couch underwent psychic transformations as he accessed his unconscious past. Freud believed unconscious thoughts could be made conscious through spoken language. During free association, the patient's spoken words permitted his unconscious thoughts to be unearthed and transformed into clear mental images. The words exiting the patient's mouth may be compared with the words and spirit leaving the mummy's mouth. Inside each sealed, active container, the resting body's internal psychic reality was externalized through spoken language.

In ancient Egypt, the opening of the mummy's mouth was a prerequisite for the deceased's "passage to the beyond-for he must use his mouth to pronounce the names of the door-guardians and answer their questions, and he must speak at the judgment to prove himself worthy to enter the afterlife" (Taylor, 2010, p. 88). After the body was laid into the sarcophagus, a stone was traditionally placed on the mummy's mouth. The stone enabled the mummy to speak during the judgment procedure and in the Land of the Dead. As Freud pointed out in his writings on the archeological metaphor: "Saxa loquuntur ['Stones talk!']" (Freud, 1896, p. 192).

Sealing the tomb permitted the mummy's spirit to transcend the body and cross over into the Land of the Dead. In Freud's office, the patient's accessing of his unconscious permitted his metaphorical crossing over into the afterlife. The enclosed active rooms made it possible for the resting body to undergo psychic transformations. Each room's central, open portal signified the space's transformative potential. The patient's and mummy's psychic transformations connoted a rebirth within the room.

FREUD'S TOMB-LIKE OFFICE ELEMENTS

The ancient Egyptian tomb architecture and the precise composition of the internal tomb elements initiated the mummy's rebirth and transition into the afterlife. The inner realms of ancient Egyptian tombs included terracotta statuettes much like those on Freud's desk. The Egyptian tomb murals painted on the durable, stone walls resembled Freud's colorful, Egyptian paintings and stone plaques. The tomb mirrors strategically positioned near the mummy's head mirrored the placement of Freud's Etruscan mirrors. Freud's couch found its double in the Egyptian sarcophagus. The tomb's false door was comparable with Freud's discrete, back exit door. Freud's office and the Egyptian tomb were filled with vast amounts of personal treasures. Essential tomb artifacts, such as canopic jars, boxes, and mummy cloth, were also present within Freud's tomb-like office. Freud's oriental rugs may be seen as analogous to traditional linen mummy cloths. The office's rugs and the tomb's mummy cloths securely held and preserved the resting mummy's and patient's bodies. By taking a closer look at some of these elements, we find how each contributes to the room's status as an active container.

RUGS AS MUMMY CLOTH

In Freud's office, soft, Persian rugs were draped over and above the couch. The ornately patterned cloths enveloped and held Freud's patients during psychoanalysis sessions. In a description of Freud's office space, Donald W. Winnicott emphasizes the comfort generated by the patient's designated space in the haptic consulting room. Winnicott observes: "the patient would be lying on a couch, that is to say comfortable . . . and probably a rug

and some water would be available" (Winnicott, 1955, p. 285). In architecture, the haptic realm is "defined by sense of touch. When the materiality of the details forming an architectural space become evident, the haptic realm is opened up. Sensory experience is intensified; psychological dimensions are engaged" (Holl, 1996, p. 16). In Freud's office, the tactile elements played an active role in opening the patient's unconscious mind. The rugs stimulated the patient's senses and psyche. The mummy laid bare his soul once protected by the haptic, linen cloths. Freud's interest in Egyptian burial practices was reflected by the linen mummy cloths displayed in his office.

Freud showcased mummy cloth "made of flaxen cloth or pure linen" in his tomb-like office (Caminos, 1992, p. 338). The displayed cloth's "21 warp threads and 11 woof threads to the centimeter" are in "agreement with the usual texture of Egyptian mummy linens" (Caminos, 1992, p. 338). These Egyptian mummy bandages were purchased by Freud around 1896 when he began collecting artworks and antiquities.

Although the year 1896 demarcates the death of Freud's father, another pivotal date looms in the background. Freud was preparing for his own death when he began collecting Egyptian antiquities. On June 22, 1894, Freud wrote to his friend and colleague Wilhelm Fliess, "I shall go on suffering from various complaints for another 4 to 8 years with good and bad periods and then between forty and fifty perish very abruptly from a rupture of the heart" (Hunter, 1989, p. 99). Freud did not die during the predicted time frame. However, Freud's father passed away.

Diana Fuss and Joel Sanders point out that in 1886 Freud's father "fell fatally ill and died of heart failure . . . Freud apparently felt that his father died in his place" (Fuss and Sanders, 2015, p. 4). Fuss and Sanders claim that Freud's preoccupation with his own death and his father's death played a central role in his "death deliria" (Fuss and Sanders, 2015, p. 3). Fuss and Sanders write that "the lingering death of Jakob Freud is generally recognized as the emotional crisis that galvanized Freud's compensatory interest in collecting" antiquities (Fuss and Sanders, 2015, p. 3).

Adding complexity to this generally accepted view, it seems relevant to point out that Freud's interest in collecting ancient Egyptian funerary antiquities preceded his father's death. In the years that Freud began dedicating himself to collecting tomb antiquities, he frequented the Louvre's ancient Egyptian collection (1985), coined the term psychoanalysis (1986), wrote "The Aetiology of Hysteria" (1986), and began writing *On the Interpretation of Dreams* (1986). Each of these endeavors is haunted by Freud's interest in ancient Egypt.

FREUD'S EARLY EGYPTIAN ENCOUNTERS

In 1885, while he was a student in Paris under Dr. Jean-Martin Charcot, Freud developed his first thoughts on what would evolve into psychoanalysis. Charcot's lectures on hypnosis and hysteria in confluence with Freud's visits to the Louvre were formative in his early conceptualizations of psychoanalysis. A year later, in 1886, Freud coined the term psychoanalysis "within the framework of a new theoretical understanding of the . . . etiology of hysteria" (Roazen, 1997, p. 39). Freud's essay "The Aetiology of Hysteria" seems to reference ancient Egyptian tomb architecture in the section that describes an archeologist's discovery of an ancient "treasure-house" bearing "numerous inscriptions," which "yield undreamed-of-information" once "deciphered and translated" (Freud, 1896, p. 192). It is known that this particular text was "influenced by the school of Charcot" (Roazen, 1997, p. 39). Much of Charcot's practice may be traced back to ancient Egypt.

Here, it is important to note that hypnosis (as performed by Charcot) was first practiced in ancient Egyptian sleep temples under the influence of Imhotep. Hysteria may also be traced back to ancient Egypt. In his office, Freud paid homage to the fathers of hypnosis. He showcased multiple Egyptian Imhotep statuettes and a reproduction of Pierre Albert-Brouillet's engraving "La Leçon Clinique du Dr. Charcot." Freud was deeply inspired by his encounters with Charcot and the ancient Egyptian ruins displayed in Paris.

In Paris, Freud frequented the Louvre's Egyptian collection. On October 19, 1885, following his first visit to the Louvre, Freud wrote to his fiancée Martha Bernays: "I just had time for a fleeting glance at the Assyrian and Egyptian rooms, which I must visit again several times" (Freud, 1975, p. 173). In the letter, Freud also describes the Egyptian obelisk from Luxor displayed at Place de la Concorde. The letter reads:

> *Imagine a genuine obelisk, scribbled all over with the most beautiful birds' heads, little seated men and other hieroglyphs, at least three thousand years older than the vulgar crowd round it, built in honor of the King whose name today only a few people can read and who, but for the monument, might be forgotten . . . For me these things have more historical than aesthetic interest (Freud, 1975, p. 173).*

Freud's Paris letters reveal early traces of his passion for ancient Egyptian antiquities as do his cherished copies of Schliemann's work. Freud's passion for collecting Egyptian funerary artifacts was sparked by the above events and may have been further amplified subsequently by his father's death. Freud's interest in ancient Egypt before his father's death and heightened interest in the time of sorrow and dread after his father's death seem to have both contributed to his creation of a tomb-like office space.

It is also useful to point out that in his letter, Freud emphasizes the "beautiful birds' heads" (Freud, 1975, p. 173). In Egyptian tomb hieroglyphs, the direction the sparrow hawk faced indicated in which direction the text should be read. Freud's later writings disclosed his awareness of this system. In *On the Interpretation of Dreams*, Freud wrote: "of sparrow hawks from an Egyptian tomb-relief" (Freud, 1900, p. 183). He noted: "to be able to read [Egyptian hieroglyphs] we . . . must depend upon the faces of the . . . birds" (Freud, 1900, p. 184).

Freud was particularly fascinated by the Egyptian writing system's "vertical rows . . . considerations of beauty and proportion" (Freud, 1900, p. 184). The composition and organization of Freud's office furniture and artifacts appear to also adhere to an organized system.

ARCHITECTURE, SPACE, AND COMPOSITION

Freud was known to arrange his antiquities thematically in his case-like shelves and on his tables. In *The Psychopathology of Everyday Life*, Freud remarks, "shortage of space in my study has often forced me to handle a number of pottery and stone antiquities (of which I have a small collection) in the most uncomfortable positions, so that onlookers have expressed anxiety that I should knock something down and break it. That however, has never happened" (Freud, 1901, p. 55). Although crowded, the object arrangements in Freud's tomb-like office followed a specific order that he knew well. Danze notes that "the position of the furniture . . . and the space between objects in the [active] room, establishes the sense of intimacy, familiarity, and safety" (Danze, 2005, p. 114).

In ancient Egypt, the spatial composition and orientation of tomb architecture, art, and furniture followed a specific scheme based on the greater environment. German Egyptologist Alexandra Verbovsek describes the basis of Egyptian architectural orientation and composition in the following excerpt:

> *[Egyptian] architecture or objects featured with images could–in a broader context–be integrated into comprehensive landscape conceptions which provide specific ways of viewing or axes of perception . . . For example, the alleys of sphinxes were visibly associated with particular buildings, assigned cult axes, and aligned the way toward temple entrances . . . guided actions and gave orientation within rooms (Verbovsek, 2014, p. 144).*

In Freud's office, furniture and object composition served as a means of orientation. The architectural scheme appeared to create telling axis lines within the space. The axis lines revealed what may be considered the room's doubles.

If we draw a diagonal line through the central, open doorway across Freud's room, we may connect the couch to his desk. We may also connect Freud's consulting room chair to his anthropomorphic study chair. If Freud directed his gaze diagonally across the room into his office, his study chair was directly in his line of vision (see **Figure 3**). Perceptual lines and strategic object orientation defined Freud's and the Egyptian's architectural work. The perceptual axis lines generated by the positioning of Freud's consulting room couch and chair produced the motif of the double. The doubles found in Freud's office contributed to the sense of security, familiarity, and safety inherent to the active room.

COUCH AS ACTIVE CONTAINER

In Freud's office, the patient's initial disorientation was countered by the comfort provided by the haptic rugs, symmetric room composition, and couch (see **Figure 4**). Danze describes the patient's experience on the couch in detail:

> *Once the analysand is invited to lie on the couch, in the reclining position, disorientation, reconfiguration and instability all naturally and fundamentally ensue. The shift from being grounded, upright, mobile and physically in control, from being in command of one's physical location, to being passive, at rest . . . is one of the most subtle*

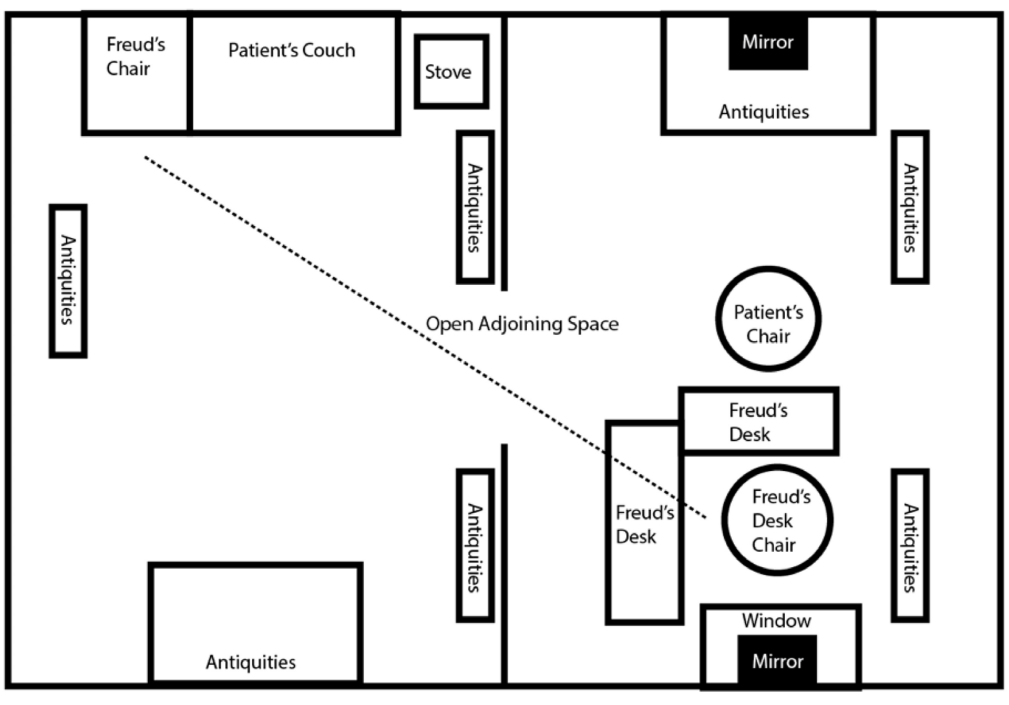

FIGURE 3 | Depiction of Freud's Office Architecture: Study **(left)** and Consulting Room **(right)**. If Freud directed his gaze diagonally across the room into his study from his consulting room chair, his study chair was directly in his line of vision (**Figure 3** by Julia K. Schroeder).

FIGURE 4 | In Freud's office, the patient's initial disorientation was countered by the comfort provided by the haptic rugs, symmetric room composition, and couch. (**Figure 4** by Edmund Engelman, courtesy of Thomas Engelman).

and powerful physical adjustments that initiates transformation (Danze, 2005, p. 113).

The haptic experience facilitated the patient's acclimation to his new position on the couch. As previously alluded to, Freud's couch symbolized the Egyptian sarcophagus. In 18th Dynasty tombs, the sarcophagus was traditionally positioned against the burial chamber wall. Freud's couch was also positioned against the wall. Freud's couch and the sarcophagus held and safeguarded the shrouded bodies.

Freud's couch and the coffin denote boxes–active containers–that held and preserved the resting body as the inner being embarked on a psychic, spiritual journey into the afterlife. According to Freud, as the patient lay on the couch, "the 'upward drive' of the repressed stirred into activity" (Freud, 1937, p. 267). The analyst's couch and its placement within Freud's office contributed to the room's status as an active agent within the psychoanalysis session. Reflecting on Freud's psychoanalysis techniques, Danze concludes that the patient's "new position and point of view [on the couch] transform the consulting room from a passive spectator to a full participant of the analytic work or from a neutral film to a receptive container of actions and movements" (Danze, 2005, p. 113). The furniture's architecture contributed to the room's initiation of the analysand's and analyst's psychic journey.

FREUD'S CHAIR

During psychoanalysis sessions, Freud also embarked on a psychic journey. Emulating the patient, he assumed a passive position as he sat in his large, comfortable, reclining chair situated behind the couch (see **Figure 5**). The spatial configuration embodies the active container's primary facet: the patient "lies down and the analyst sits behind him" (Quinodoz, 1992, p. 630). Seated in the room's corner, Freud was outside of the patient's line of vision. Here, Freud leaned back and assumed a resting

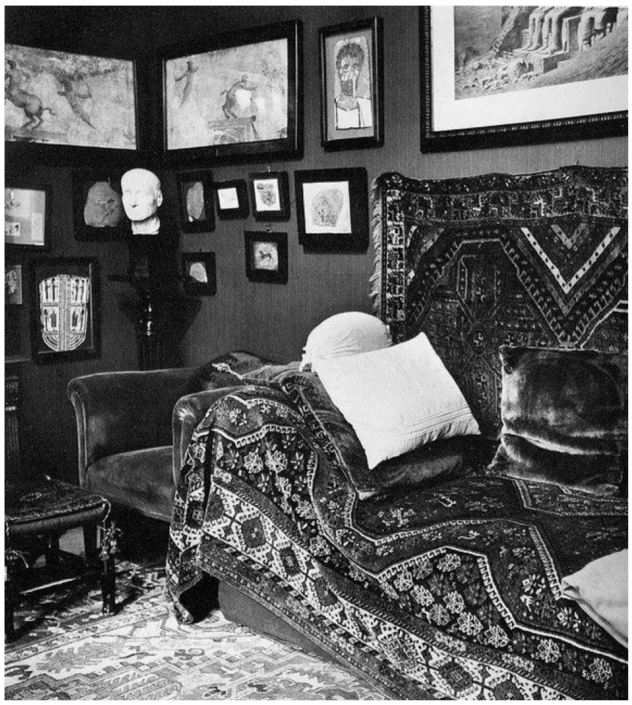

FIGURE 5 | Freud assumed a passive position as he sat in his large, comfortable, reclining chair situated behind the couch. (**Figure 5** by Edmund Engelman, courtesy of Thomas Engelman).

position. While the room remained active, the analyst played a passive role during psychoanalysis sessions. During sessions, Freud practiced a passive technique, which he termed "evenly-suspended attention."

In "Recommendations to Physicians Practicing Psychoanalysis," Freud describes the technique and his role within the contained office space. He recommends the analyst "should simply listen, and not bother about whether he is keeping anything in mind" (Freud, 1912, p. 77). Danze points out that "when the analysand is looking up, out, or over, and looking not at any particular focal point, there is made to exist a spatial openness, an implied spatial potential for infinity" (Danze, 2005, p. 112). Both chair and couch permitted the bodies to lean back comfortably and submit themselves to the forces of their unconscious.

ACTIVE ROOM AS WOMB

The couch, coffin, and chair may also be interpreted as active containers. These vessels shelter and contain the body as it undergoes transformative processes. Each contained space was an "envelope and backdrop for what goes on in it, a sensitive vessel that holds all actions and movements, containing and holding the patient" (Danze, 2005, pp. 113–114). The office itself represented a maternal, comforting space. Winnicott's description of the psychoanalyst's need to create a maternal "holding environment" resonates with Freud's clinical setting (Winnicott, 1965, p. 239).

For Freud, the "dwelling-house was a substitute for the mother's womb, the first lodging . . . in which he was safe and at ease" (Freud, 1930, p. 152). Herein lies a semiotic similarity between the languages of architecture and psychoanalysis. Freud's office appears analogous to the womb. The office matured the patient until he was ready to move on from the maternal holding space.

Inside the womb and the tomb, the bodies were subject to forces beyond their control as they underwent formative transformations. Each holding vessel evokes the active container and active room frameworks. According to Danze's "active room" theory, "the room itself has the power to initiate and sustain a shift in the analysand . . . The analytic space is then both a natural witness and powerful participant in the psychoanalytic relationship" (Danze, 2005, p. 110).

WALLS

In the active room, the walls create an enclosed space and provide a thinking interface for the patient. According to psychoanalyst Julie Leavitt, upon the interface "echoes of memories, the history, and the transference interact with the very physical dimension and make live and use again the visible features of the analytic room" (Schinaia, 2018, p. 181). The walls enclosing Freud's office space created an interface for memories to play out on. The patient conjured up associations and connections as he gazed at Freud's walls. Significantly, a large color print of Ramse's temple at Abu Simbel with its colossal seated deities emerging out of a rock wall was located on the wall directly above the couch. Various memories full of hidden meanings would arise in the patient, who was positioned facing this wall. Freud's patient Hilda Doolittle (H.D.) described the print in her *Tribute to Freud*:

> *In one of our talks in the old room at Berggasse, we had gone on one of our journeys. Sometimes the Professor knew actually my terrain, sometimes it was implicit in a statue or picture, like that old-fashioned steel engraving of the Temple at Karnak [the Ramse's temple print] that hung above the couch. I had visited the particular temple, he had not (Doolittle, 1956, p. 9).*

In a sense, the walls served as a memory substrate, reminiscent of Freud's descriptions of the "mystic writing pad." The psyche projected itself onto the physical environment.

In Egypt, "so much of the daily life [was] to be represented on the walls of the tomb chapel" (Smith and Simpson, 1999, p. 1). The walls preserved the mummy's past life and memories, which would accompany him into the afterlife. To ensure that the past was preserved on the tomb walls, in ancient Egypt "enduring stone forms were chiefly reserved for the building of temples and tombs" (Smith and Simpson, 1999, p. 1). The strong stone walls "lead to the excellent preservation of so much which represents the peculiarly Egyptian emphasis upon tomb architecture" (Smith and Simpson, 1999, p. 1). The walls in Freud's office and the Egyptian tomb functioned as a means of memory activation and preservation. The active room's walls had the "capacity to evoke different kinds of associations" (Danze, 2005, p. 123). The psyche may be interpreted as having been preserved within the room's physical structures, including the walls and terra cotta statues.

TERRA COTTA STATUES

The Egyptians believed that after death the body was transfigured into the "ka," an emanation of the spirit. The mummy's ka had a separate, spiritual existence from the body and was described as the mummy's "double or as a protective genius" (Smith and Simpson, 1999, p. 5). Resonating with Freud's unconscious–conscious dualism, the Egyptians believed that human beings possessed two bodies–a spiritual one and a physical one. The terra cotta tomb statues served as the physical double of the ka. The terra cotta statues were traditionally placed near the sarcophagus. If the body was stolen, the statues stood in for the body.

André Bazin describes the significance of the Egyptian statues' placement in detail: "the Egyptians placed terra cotta statuettes, as substitute mummies . . . It is this religious use, then, that lays bare the primordial function of statuary, namely, the preservation of life by a representation of life" (Bazin, 1960, p. 5). Bazin terms the immortality made possible by the mummy's terra cotta double the "mummy complex." Similarly, in "The Uncanny," Freud claimed the ancient Egyptian "'immortal' soul was the first 'double' of the body" (Freud, 1919, p. 9). Freud asserted that "this invention of doubling as a preservation against extinction has its counterpart in . . . the Ancient Egyptians . . . art of making images of the dead in lasting materials" (Freud, 1919, p. 9). We may read this as a reference to the terra cotta statues created by the ancient Egyptians.

In Freud's office study, Egyptian terracotta statues were lined up across his desk (see **Figure 6**). In his lectures, Freud describes his desk's topographical arrangement as "a circle of bronze statuettes with small terra-cotta figures . . . set behind this inkstand" (Freud, 1901, p. 54). Freud emphasizes the statuette's strategic placement by stressing their designated space "behind" the inkstand. A diagonal axis line drawn between the study and consulting room would appear to connect the statuettes standing atop Freud's desk to the patient atop Freud's consulting room couch. The statues appear to symbolize doubles of the patient. The statues' positioning recalls that of the statues found near the sarcophagus and mummy in the ancient Egyptian tomb.

The statues' positioning within the active room could be interpreted as permitting the patient's doubling and rebirth. In other words, the statues served as the patient's double as he was reborn during the psychoanalysis session. The statues aided Freud's patients as they embarked upon their spiritual journey into the next enlightened life. These figurines played an active role within psychoanalysis sessions. In Charles Rice's study entitled "Lost Objects: Sigmund Freud's Psychoanalytical Interior," the active role of Freud's statues is also investigated (Rice, 2006, p. 46). Rice writes the antiquities may be seen as "elements of a patient's psychological history" (Rice, 2006, p. 46). He also writes that the antiquities speak to the analyst, who must translate what is said into something meaningful. Rice's claims further supplement the argument that Freud's statues represent the patient's double.

The active participation of the statues and their role as the patient's double were also emphasized by Freud's tendency to interact with his statues during psychoanalysis sessions. In her memoirs, Freud's patient Hilda Doolittle writes, "we looked over the images in one of the other cases . . . The Professor

FIGURE 6 | In Freud's office study, Egyptian terracotta statues were lined up across his desk. A mirror was framed by Freud's window. (**Figure 6** by Edmund Engelman, courtesy of Thomas Engelman).

brought out a wooden Osiris . . . blackened by time . . . there was another green-blue Osiris. The Professor said, 'They are called the answerers, as their doubles or ka's come when called"' (Doolittle, 1956, p. 172).

Freud associates the double not only with the Egyptian tomb's plastic arts but also with the mirror. Drawing upon Otto Rank, Freud examines "the connections the 'double' has with reflections in mirrors, with shadows, guardian spirits, with the belief in the soul and the fear of death" (Freud, 1919, p. 9). The intimate relationship between the double and the mirror may be traced back to the ancient Egyptian tomb.

MIRRORS

The mirrors in Freud's office appear to provide a further parallel to the historically Egyptian motif of the double. Historically, "the Egyptians may have believed that the mirror helped preserve the ka, the double discovered in the mirror's depths, and allowed it to make a transition to another life. Thus, mirrors are frequently depicted on the wall paintings directly before the face of the deceased" (Pendergrast, 2009, p. 5). In ancient Egypt, such mirrors were considered light reflectors that guided the way into the afterlife.

In Freud's study, one mirror was positioned directly behind his desk. This mirror was framed by Freud's window (see **Figure 6**). During initial consultation meetings, Freud's patients sat directly across from this mirror (see **Figure 7**). The mirror framed the patient's face and created his double while he sat in his designated position across from Freud.

Freud's second mirror was placed directly across from his desk. When directing his gaze straight ahead, Freud encountered his own double (see **Figure 3**). This physical setup was perhaps intentional. In "Papers on Technique," Freud claimed that "the doctor . . . should be opaque to his patients and, like a mirror, should show them nothing but what is shown to him" (Freud, 1914, p. 118). Freud often remained opaque and expressionless during consultation meetings, recalling the mummy's death mask

FIGURE 7 | During initial consultation meetings, Freud's patients sat directly across from Freud and the mirror. (**Figure 7** by Edmund Engelman, courtesy of Thomas Engelman).

in ancient Egypt. The purpose of the Egyptian death mask painted onto the sarcophagus was "to provide a face for the dead in the afterlife" (Bryant and Peck, 2009, p. 322).

THE ACTIVE ROOM AS A TOMB

Freud's writings indicate his interest in the Egyptian afterlife. In his lectures, Freud asserts:

> *Every person acquainted with ancient rite knows how seriously, for example, the Egyptians considered the portrayal of a journey to the land of the dead. There still exist many copies of the "death book" which was given to the mummy for this journey as a sort of Baedeker. Since the burial places have been separated from the living quarters, the last journey of the dead person has become a reality (Freud, 1917, p. 126).*

Here, Freud emphasizes the importance of isolating the tomb as a prerequisite for the Egyptian final journey to begin. Upon his death, Freud spent his final days in his tomb-like office, which he recreated in London at the end of his life.

In 1938, Freud fled Nazi-occupied Vienna. Freud reconstructed his tomb-like office in London with the help of his architect son, Ernst Freud. To "recreate the interior of Sigmund Freud's Viennese study and consulting room was . . . [a] goal of the remodeling of 20 Maresfield Gardens" (Welter, 2011, p. 151). In order to maintain the previous office's architectural composition, Ernst Freud "knocked together" two rooms by removing a wall (Welter, 2011, p. 151). Although Freud's office no longer exhibited a central open doorway, the space remained composed of two conjoined rooms. Similar to Berggasse 19, the new office at 20 Maresfield Gardens remained separated from the Freud family's living quarters. Fuss and Sanders note that Freud's maid Paula Fichtl arranged the London office "to reproduce, as closely as possible, the office at Berggasse 19" (Fuss and Sanders, 2015, p. 15). For the most part, Freud's office retained its previous spatial and architectural composition.

Freud only disbanded his clinical practice in London 2 months before his death. "During the final days, Freud requested that his bed be brought down to the study so that he could be near his books, desk, and beloved collection of antiquities" (Cohen, 2014, p. 3). In Freud's tomb-like office, he assumed the role of

the mummy. Freud quietly died on September 23, 1939. Freud was embalmed "in a tomb he spent over 40 years preparing" (Fuss and Sanders, 2015, p. 15). The preservation of Freud's ashes in his London office amplifies this claim. Freud's Egyptian tomb-like office immortalized his collection and legacy. Freud's active room preserved his patient's memories and his own. Drawing upon Diana Fuss and Joel Sanders, Charles Rice writes: "'The location of the analytic scene within the four walls of a crypt . . . at Freud's office . . . was an overdetermined space of loss and absence, grief and memory, elegy and mourning' . . . This . . . was the spatialization of Freud's own psychological interiority" (Rice, 2006, p. 43).

DISCUSSION AND WIDER DISCOURSE

The notion that architecture represents a spatialization of the psychological interiority is further reflected in works that posit architecture's status as an active, psychic entity. Therefore, the present study may also be seen as part of the wider discourse around the psychic potential of architecture. Contemporary architectural theorists, such as Juhani Pallasmaa and Lucy Huskinson, reference architecture's active agency and psychic potential. These theorists oppose those who state that the therapeutic space (in the physical sense) should be neutral.

Those in support of the therapy room's neutrality argue that personalized office design and architecture could "give indications of the analyst's personality and interface with the analysand's [patient's] phantasmatic and associative freedom" (Schinaia, 2018, p. 194). Expanding upon this viewpoint, psychoanalyst Christopher Bollas writes that "few physical distractions" permit the self "to recline into [the mind's] interiority" (Bollas, 2012, p. 53). Neutral therapy spaces are often characterized by "modern architecture," "reduction, simplification, and deprivation (this idea is based on the German architect Ludwig Mies van der Rohe's maxim *less is more*)" (Schinaia, 2018, p. 190). These spaces generally include "an armchair, a couch, sometimes a writing desk and closet" (Schinaia, 2018, p. 190). The tension between those who assert that the therapeutic space should be neutral and those who believe in the psychic agency of space may be considered a reaction to the "'too full spaces' of Freud's and [the] first Freudian analysts' rooms" (Schinaia, 2018, p. 190).

Today's analysts seem to have renounced the "neutrality that was recommended until a few decades ago . . . in their exterior settings" (Bolognini, 2010, p. 8). Juhani Pallasmaa "opposes the notion of (supposed) neutrality of the environment and supports the need to 're-sensualise' architecture" (Schinaia, 2018, p. 159). Pallasmaa asserts "profound architecture makes us experience ourselves as complete embodied and spiritual beings" (Pallasmaa, 2012, p. 13). Along a similar vein, Lucy Huskinson writes: "Buildings are active participants in the creation and development of personality" (Huskinson, 2018, p. 15). Citing Gaston Bachelard, whose book *The Poetics of Space* was formative in forging the relationship between psychoanalysis and architecture, Huskinson states: "'Everything comes alive when contradictions accumulate' . . . Through tensions of opposites the correspondence between psyche and place . . . is consolidated, with both aspects undergoing some sort of 'expansion' or enrichment as a result" (Huskinson, 2018, p. 135). In the office and tomb, the resting body could be viewed as undergoing a psychic expansion and enrichment. Therefore, we may situate the present analysis within the wider discourse around the active, psychic potential of architecture.

FIGURE 8 | The office space at Berggasse 19 in Vienna actively influenced Freud's practice, work, and legacy. (**Figure 8** by Edmund Engelman, courtesy of Thomas Engelman).

This study may also be productive to read in dialogue with the work of Diana Fuss and Joel Sanders. As previously stated, this study builds upon the article "Berggasse 19: Inside Freud's Office" by Diana Fuss and Joel Sanders. Although Fuss and Sanders do not put the Egyptian office elements and architecture into the foreground and illuminate the office more as Freud's own tomb, rather than a place where the patient undergoes a tomb-like transformation, their allusions to tombs strengthen the present argument. Whereas Fuss and Sanders explore "the role of both vision and hearing in three-dimensional space, examining how architecture organizes the physical and sensory interaction of bodies as they move through the interior of Freud's study and consulting room," the present study seeks to emphasize the active role the Egyptian architecture and elements played within Freud's tomb-like office (Fuss and Sanders, 2015, p. 1). Both studies ultimately seek to shed light on the significant role Freud's office architecture played in his psychoanalytic practice.

CONCLUDING REMARKS

The office space actively influenced Freud's practice, work, and legacy (see **Figure 8**). In this study, I posit and provide support for the notion that the ancient Egyptian tomb elements and architecture in Freud's office actively contributed to his psychoanalysis practice. Freud's office and the ancient Egyptian tomb represent active rooms. These closed-off spaces significantly contributed to the contained individual's psychic transformation. These active containers facilitated psychic transformations by liberating the mummy's spirit and the patient's unconscious. In our present historical moment, Freud's office design continues to

influence the architectural design of many psychoanalytic offices around the world. The psychoanalyst's office remains an active vessel that facilitates transformation.

AUTHOR CONTRIBUTIONS

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

ACKNOWLEDGMENTS

The author would like to thank Dr. Winfried Kudszus for providing insightful feedback on early drafts of the manuscript and Thomas Engelman for kindly permitting the inclusion of Edmund Engelman's photographs. The author would also like to thank the reviewers for their invaluable comments on the manuscript. JS was supported by a fellowship from the University of California, Berkeley, Department of German.

REFERENCES

Bazin, A. (1960). The ontology of the photographic image. *Film Q.* 13, 4–9. doi: 10.2307/1210183

Bollas, C. (2012). *China on the Mind*. Milton Park: Taylor & Francis.

Bolognini, S. (2010). *Secret Passages: The Theory and Technique of Interpsychic Relations*. London: Routledge.

Bryant, C., and Peck, D. (2009). *Encyclopedia of Death and the Human Experience*. Thousand Oaks, CA: SAGE Publications, Inc.

Caminos, R. A. (1992). On ancient egyptian mummy bandages. *Orientalia NOVA Ser.* 61, 337–353.

Cohen, L. (2014). *How Sigmund Freud Wanted to Die*. Washington, DC: The Atlantic.

Danze, E. A. (2005). An architect's view of introspective space the analytic vessel. *Annu. Psychoanal.* 33, 109–124.

Deleuze, G., and Guattari, F. (2004). *Anti-Oedipus*. London: Bloomsbury.

Doolittle, H. (1956). *Tribute to Freud*. New York, NY: New Directions.

Freud, S. (1896). "The aetiology of hysteria," in *The Standard Edition of the Complete Psychological Works of Sigmund Freud*, ed. J. Strachey (London: Hogarth Press).

Freud, S. (1900). "The interpretation of dreams," in *The Standard Edition of the Complete Psychological Works of Sigmund Freud*, ed. J. Strachey (London: Hogarth Press).

Freud, S. (1901). "The psychopathology of everyday life," in *The Standard Edition of the Complete Psychological Works of Sigmund Freud*, Vol. 7, ed. J. Strachey (London: Hogarth Press).

Freud, S. (1909). "Notes upon a case of obsessional neurosis (the cases of 'little hans' and the 'rat man')," in *The Standard Edition of the Complete Psychological Works of Sigmund Freud*, Vol. 10, ed. J. Strachey (London: Hogarth Press).

Freud, S. (1912). "Recommendations to physicians practicing psychoanalysis," in *Standard Edition of the Complete Psychological Works of Sigmund Freud*, Vol. 12, ed. J. Strachey (London: Hogarth Press).

Freud, S. (1914). "Papers on technique," in *The Standard Edition of the Complete Psychological Works of Sigmund Freud*, Vol. 12, ed. J. Strachey (London: Hogarth Press).

Freud, S. (1917). "Introductory lectures on psychoanalysis," in *The Standard Edition of the Complete Psychological Works of Sigmund Freud*, Vol. 16, ed. J. Strachey (London: Hogarth Press).

Freud, S. (1919). "The uncanny," in *The Standard Edition of the Complete Psychological Works of Sigmund Freud*, Vol. 17, ed. J. Strachey (London: Hogarth Press).

Freud, S. (1930). "Civilization and its discontents," in *The Standard Edition of the Complete Psychological Works of Sigmund Freud*, Vol. 21, ed. J. Strachey (London: Hogarth Press).

Freud, S. (1931). "Female sexuality," in *The Standard Edition of the Complete Psychological Works of Sigmund Freud*, Vol. 21, ed. J. Strachey (London: Hogarth Press).

Freud, S. (1937). "Constructions in analysis," in *The Standard Edition of the Complete Psychological Works of Sigmund Freud*, Vol. 23, ed. J. Strachey (London: Hogarth Press).

Freud, S. (1938). "Findings, ideas, problems," in *The Standard Edition of the Complete Psychological Works of Sigmund Freud*, Vol. 23, ed. J. Strachey (London: Hogarth Press).

Freud, S. (1975). *Letters of Sigmund Freud*. New York, NY: Basic.

Freud Museum London (2018). *Highlights of Freud's Library*. London: Freud Museum London.

Fuss, D., and Sanders, J. (2015). *Berggasse 19: Inside Freud's Office (with Diana Fuss)*. Avaialble at: http://joelsandersarchitect.com/berggasse-19-inside-freuds-office-with-diana-fuss/ (acccessed April 10, 2020).

Gardiner, M. (1973). *The Wolf-Man and Sigmund Freud*. London: Penguin.

Holl, S. (1996). *Intertwining*. New York, NY: Princeton Architectural.

Hunter, D. (1989). *Seduction and Theory: Readings of Gender, Representation, and Rhetoric*. Champaign, IL: University of Illinois.

Huskinson, L. (2018). *Architecture and the Mimetic Self A Psychoanalytic Study of How Buildings Make and Break Our Lives*. Milton Park: Taylor & Francis.

Pallasmaa, J. (2012). *The Eyes of the Skin: Architecture and the Senses*. Hoboken, NJ: John Wiley & Sons.

Pendergrast, M. (2009). *Mirror, Mirror: A History Of The Human Love Affair With Reflection*. New York, NY: Basic Books.

Quinodoz, D. (1992). The psychoanalytic setting as the instrument of the container function. *Int. J. Psycho Anal.* 73, 627–635.

Rice, C. (2006). *The Emergence of the Interior: Architecture, Modernity, Domesticity*. Abingdon: Routledge.

Roazen, P. (1997). *Freud Under Analysis: History, Theory, Practice: Essays in Honor of Paul Roazen*. New Jersey: Jason Aronson.

Schinaia, C. (2018). *Psychoanalysis and Architecture: The Inside and the Outside*. Abingdon: Routledge.

Schliemann, H. (1878). *Mycenae: A Narrative of Researches and Discoveries at Mycenae and Tiryns*. Lancaster: Armstrong and Company.

Smith, W. S., and Simpson, W. K. (1999). *The Art and Architecture of Ancient Egypt*. Yale, NY: Yale University Press.

Taylor, J. H. (2010). *Journey Through the Afterlife: Ancient Egyptian Book of the Dead*. Cambridge, MA: Harvard University Press.

Verbovsek, A. (2014). *Reception and Perception. In A Companion to Ancient Egyptian Art*. Hoboken, NJ: John Wiley & Sons.

Welter, V. (2011). *Ernst L. Freud, Architect: The Case of the Modern Bourgeois Home*. New York, NY: Berghahn Books.

Winnicott, D. W. (1955). Metapsychological and clinical aspects of regression within the psycho-analytical set-up. *Paper Presemted at the Pediatrics to Psychoanalysis: Collected Papers*, (New York, NY: Basic).

Winnicott, D. W. (1965). *The Maturational Processes and the Facilitating Environment*. London: The Hogarth Press.

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 in Psychology

BRIEF RESEARCH REPORT
published: 04 August 2020
doi: 10.3389/fpsyg.2020.01861

How Does Embodying a Transgender Narrative Influence Social Bias? An Explorative Study in an Artistic Context

Marte Roel Lesur[1,2], Sonia Lyn[1] and Bigna Lenggenhager[1]**

[1] *Department of Psychology, Cognitive Neuropsychology With Focus on Body, Self, and Plasticity, University of Zurich, Zurich, Switzerland,* [2] *BeAnotherLab, Barcelona, Spain*

Virtual reality (VR) protocols inducing illusory embodiment of avatars have shown a positive impact on participants' perception of outgroup members, in line with the idea that the simulation of another's sensorimotor states might underlie prosocial behavior. These studies, however, have been mostly confined to laboratory settings with student populations and the use of artificial avatars. In an interdisciplinary effort benefiting from the heterogeneous sample within a museum, we aimed at quantifying changes in interpersonal perception induced by embodying a transgender man narrating his life. We compared an artistic methodology mixing VR and elaborate sensorimotor stimulation to a more conventional primarily audiovisual VR experience. We tested how these affect embodiment and the perception of transgender men as measured by a brief implicit association test and a questionnaire. Neither significant difference in embodiment nor changes in implicit or explicit bias was found, the latter potentially due to the initially low bias in the group. We further assessed participants' illusory embodiment as a function of age, finding a negative correlation between these. The results are discussed with respect to current theories of embodiment, differences between laboratory and real-life settings, and the intersection of art and science.

Keywords: embodiment, social cognition, naturalistic settings, art–science dialogue, sensorimotor sharing

OPEN ACCESS

Edited by:
David Harris Smith, McMaster University, Canada

Reviewed by:
Tabitha C. Peck, Davidson College, United States
Matthieu M. de Wit, Muhlenberg College, United States

****Correspondence:***
Marte Roel Lesur
marteroel@gmail.com
Bigna Lenggenhager
bigna.lenggenhager@gmail.com

Specialty section:
This article was submitted to Theoretical and Philosophical Psychology, a section of the journal Frontiers in Psychology

Received: *29 January 2020*
Accepted: *07 July 2020*
Published: *04 August 2020*

Citation:
Roel Lesur M, Lyn S and Lenggenhager B (2020) How Does Embodying a Transgender Narrative Influence Social Bias? An Explorative Study in an Artistic Context. Front. Psychol. 11:1861. doi: 10.3389/fpsyg.2020.01861

INTRODUCTION

Human beings are individuals who are consciously aware of their selves and whose perception, cognition, and behavior are grounded in bodily processes (e.g., Dijkerman and Lenggenhager, 2018). Humans perceive themselves as a distinct physical entity, and experience the world from an embodied first-person perspective (Damasio, 1994). Yet, as a highly social species, humans can simulate and mimic other persons' bodily states (Bernhardt and Singer, 2012); a process that arguably underlies social cognition, empathy, and prosocial behavior and thus might be core to society (Decety, 2010). Research has demonstrated that sensorimotor sharing [i.e., the interpersonal synchrony of the sensorimotor system at the phenomenal, functional, or neural level (Gallese and Sinigaglia, 2011; de Waal and Preston, 2017; Lamm et al., 2019)] and interpersonal perception are bidirectionally linked, such that humans tend to share more sensorimotor states with people they like and like people more with whom they share such states (van Baaren et al., 2009).

Correspondingly, sensorimotor sharing is generally reduced for outgroup as compared to ingroup members (e.g., Avenanti et al., 2010). Recent work has argued that boosting sensorimotor sharing by exposing participants to a multisensory first-person perspective narrative/experience of an outgroup member might enhance empathy toward them (e.g., Galli et al., 2015; Maister et al., 2015). This process seems to rely on altering embodiment, the somatic awareness of the body, which is claimed to be grounded in the senses of body ownership and agency (Gallagher, 2000; Longo et al., 2008). Virtual reality (VR) techniques facilitate illusory embodiment through another's first-person perspective together with the feeling of agency for and ownership of that person's body (Petkova and Ehrsson, 2008; Roel Lesur et al., 2018), presumably maximizing sensorimotor sharing. Indeed, embodiment of a virtual outgroup member has shown to decrease implicit bias toward them (Peck et al., 2013; Maister et al., 2015; Banakou et al., 2016). Real-life scenarios in diverse samples would seem particularly relevant in social cognition investigations given the study matter. However, so far in behavioral research in general (Shamay-Tsoory and Mendelsohn, 2019), as well as in studies linking embodiment and social behavior (e.g., Peck et al., 2013; Maister et al., 2015; Banakou et al., 2016), these experiments have been mostly conducted in the laboratory, with a limited population of certain age range and socioeconomic background, which might bias the results (Henrich et al., 2010). In the current study, VR techniques were used, which have been argued to be beneficial to psychological research, because of their capacity in reproducing controlled stimulation settings of high ecological validity (Pan and Hamilton, 2018) outside the confined limits of a laboratory (Mottelson and Hornbæk, 2017). While artistic groups are applying VR methods derived from psychological research on the relation of embodiment and interpersonal perception in diverse settings across cultural boundaries (e.g., Bertrand et al., 2014, 2018; Sutherland, 2015), this study aimed at using these artistic applications to feedback to science. An artistic project using video-based VR with scientific measures from laboratory settings were combined to investigate in the context of a museum exhibition how social bias toward the transgender community might change after embodying a transgender man while listening to his personal story. While arguably a higher degree of ecological sensorimotor interactions from another's perspective would maximize sensorimotor sharing, there seems to be no confirmatory evidence that more sensorimotor contingencies would result in a stronger reduction of bias compared to conventional (primarily audiovisual) VR. This comparison, however, could be important for potential applications of VR. The participants' explicit and implicit bias toward transgender men was assessed before and after the exposure to an embodied, first-person perspective immersive video narrative (part of the Library of Ourselves; BeAnotherLab, 2020) shown on a head-mounted display (HMD). This experience was created by a group professionally working on enhancing empathy and prosocial behavior with focus on artistic methods (Sutherland, 2015; Bertrand et al., 2018). In two different groups, the influence of distinct degrees of multimodal synchrony of this experience on interpersonal perception was assessed. The audiovisual content was the same in both, but the presentation was either *conventional*, involving only vision, sound, and free head movements; or *sensorimotor*, that is, additionally involving enhanced somatosensory–motor interactions (i.e., synchronous tactile, motor, and proprioceptive signals). In the narrative, a transgender man talks about his life and journey of transformation. Building on the broad literature on multisensory stimulation and embodiment (e.g., Ehrsson, 2012) showing an important effect of multisensory coherences for illusory embodiment (Botvinick and Cohen, 1998; Kalckert and Ehrsson, 2012; Roel Lesur et al., 2018), a stronger embodiment illusion in the *sensorimotor* compared to the *conventional* group was hypothesized (hypothesis 1). Because experiencing an outgroup member's story from a first-person perspective has shown to reduce bias toward the outgroup (Galli et al., 2015), and the content of the video-based experience has been developed to enhance empathy, reduced postexperience implicit and explicit biases in both the *conventional* and the *sensorimotor* group were expected, as compared to a *control* group, which had been subjected to an unrelated VR experience (hypothesis 2). Given the link between sensorimotor sharing with outgroup members and prosocial conducts toward them (e.g., Maister et al., 2015), a stronger reduction of bias toward the *sensorimotor* as compared to the *conventional* group was expected (hypothesis 3). Lastly, following evidence pointing at bodily self-plasticity in adults decreasing with age (Tajadura-Jiménez et al., 2012; Graham et al., 2015) and given the anticipated broad age range in the museum sample, a negative correlation between illusory embodiment with age was hypothesized (hypothesis 4).

MATERIALS AND METHODS

Participants

A total of 71 individuals [age ranging from 20 to 72 years (mean, 34.1; SD = 13.3); 41 females] were recruited at a museum within the exhibition "100 Ways of Thinking" at the Kunsthalle Zurich[1]. They were assigned to a group with enhanced sensorimotor (i.e., involving touch, movement, and interaction with objects) stimulation [*sensorimotor* group, n = 49; mean age = 33.5 (SD = 12.0); 29 women] or to a group with conventional VR relying on audiovisual stimulation only [*conventional* group, n = 22; mean age = 35.3 (SD = 16.0); 12 women]. Both groups had free and visually contingent head movements. With the intention to record the maximum number of participants, more people were assigned to the *sensorimotor* than the *conventional* group because the shown art project was initially designed for such setting, and attendees showed more interest in it. All protocols were approved by the ethics committee of the Faculty of Arts and Social Sciences at the University of Zurich (approval no. 17.12.10). A total of 43 additional individuals (*control* group) were recruited at the University of Zurich in the framework of another study using a similar body illusion, which was unrelated to the theme of gender identity (unpublished work; the setup was comparable to Roel Lesur et al., 2020) as a control

[1] http://kunsthallezurich.ch/en/100-ways-thinking

group [mean age = 22.40 (SD = 3.1); 32 women]. A Mann–Whitney comparison showed a significant difference of age between participants recorded at the museum and the laboratory ($U = 2654$, $p < 0.001$). Participants gave written informed consent to participate in the study, which was approved by the ethics committee of the Faculty of Arts and Social Sciences at the University of Zurich (approval no. 17.12.15).

Materials

An Oculus CV1 HMD was used for stimulation. The software was designed using Unity 2017 for displaying a 235-degree prerecorded video portraying the first-person perspective as real person. It ran on a PC (Nvidia Geforce GTX 1080 8 GB; 16 GB RAM; Intel Core i7, 3.2 GHz). The video (see **Supplementary Material**) was filmed from the perspective of the transgender man using a Kodak SP360 4K camera at a resolution of 2,160 × 2,160 pixels at 30 fps; synchronous binaural audio recording of the environment was recorded from his ears using a Zoom H4 recorder and Roland CS-10EM microphones. Additional audio of his voice was recorded separately and superimposed to the environmental sounds. The questionnaires were displayed and answered using a computer monitor and a mouse; the tasks were programmed with JavaScript. Audio cues for the researchers appeared on headphones and were synchronized with the video.

General Procedure

See **Figure 1** for the general procedure in the three different groups.

VR Experiences

Gender Identity Narrative Used at the Museum

The stimulation procedure at the museum consisted of displaying on an HMD an audiovisual narrative about Jonah, a transgender man talking about his life, friends, identity, relationships, and body transformation. Jonah's story is told by himself and accompanied by different audiovisual environments of his hometown and close people shown from his perspective. The content duration is 7 min 47 s (see **Supplementary Material** for the content).

For the *sensorimotor* group, the feeling of the wind and interactions with objects and other people were reenacted by the experimenter; Jonah's movements were reenacted by the participant. For example, when participants saw somebody touching their shoulder, they were touched by the experimenter to achieve multimodal synchrony (**Figure 1**), which has been described to induce a stronger sense of embodiment (Petkova and Ehrsson, 2008; Maselli and Slater, 2013; Roel Lesur et al., 2018). Stimulation was administered as follows: wind by manually fanning with cardboard, human touch with the experimenter's hands, and self-touch by having participants touch themselves. Interactions with objects were facilitated as follows: using a chair to sit down, a smartphone to grab and touch, and a piece of cardboard (approximately corresponding to the size of a skateboard) to move with a foot. A reference to how these interactions were performed in another setting can be found in KCTS9 (2018). The experimenters wore headphones where instructions and timing for each action were presented to synchronize with the VR content. Additionally, they could visually monitor what participants saw in the HMD. Previous extensive rehearsal of the actions was performed.

The *conventional* group did not take part in the multimodal interaction, but sat down and experienced the audiovisual content presented in the HMD. They could still move their heads to explore the visual environment.

Irrelevant Body Illusion Used in the Laboratory Setting

For the *control* group, elements of this study were integrated into another embodiment-related experiment (similar to Roel Lesur et al., 2020) conducted at the University of Zurich in a laboratory setting. This group was included with the aim of disentangling the potential effect of retesting on both measures of bias. Participants in the *control* group were, similarly, subjected to a body illusion involving multimodal stimulation with the same HMD model; crucially, they were not exposed to any themes of gender identity or any narrative-based stimulation. The stimulation lasted approximately 30 min.

Measures

Note that in order to prevent participant bias on the questionnaires due to the public setting in the museum, only they would look at the screen during the procedure and potential interruptions would be prevented by the experimenters.

Explicit Attitudes Toward Transgender

Before and after stimulation, participants were required to answer on a visual analog scale (VAS) presented on a PC an *Attitudes Toward Transgender Men and Women* questionnaire (Billard, 2018). Six of 12 items were selected that were answered on a scale ranging from strongly disagree (0) to strongly agree (100). The items were "transgender men. . . (1) will never really be men, (2) are only able to look like men but not be men, (3) are unable to accept who they really are, (4) are trying to be someone they are not, (5) are denying their DNA, (6) are unnatural."

Brief Implicit Association Test

Participants' implicit attitudes were assessed before and after exposure using a good-focal Brief Implicit Association Test (BIAT; Sriram and Greenwald, 2009) involving four categories of stimuli: a pair of contrasting attributes (positive vs. negative) and a pair of target images depicting a change in haircut of men (biological men) or a transition from a woman to man (transgender men; **Figure 1**). Note that while the pictures depict transgender women due to the availability of the images (Gonzalez, 2015), they included an arrow showing the transition in the other direction. The BIAT is a validated brief version of the well-known Implicit Association Test (Greenwald et al., 1998). Before the task, participants were given time to familiarize themselves with the images and their respective categories. The task was presented on a PC, and keystrokes were performed on a keyboard.

The BIAT consisted of six blocks in total, of which only two served as experimental blocks. In congruent blocks, participants

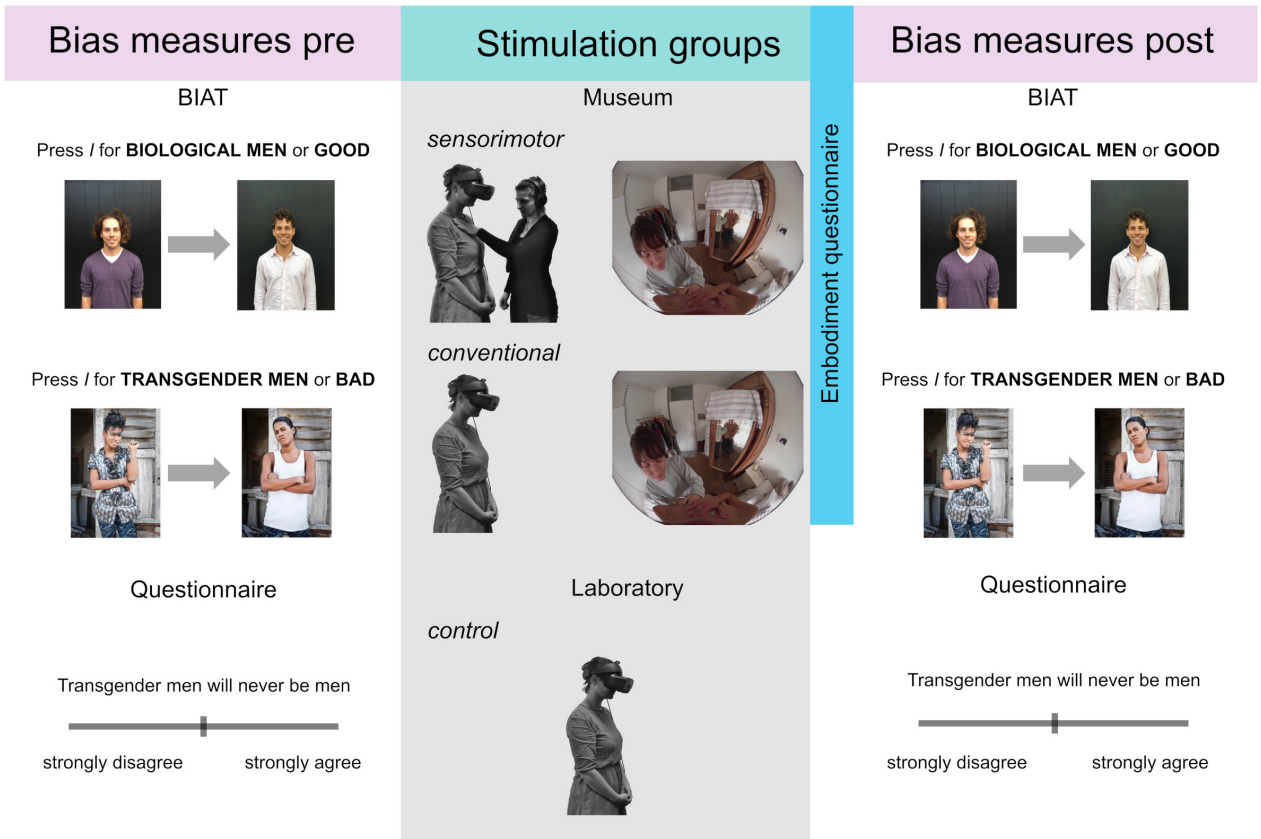

FIGURE 1 | Experimental procedure showing the implicit and explicit bias tasks and groups differing by intervention. The *control* group was subject to an irrelevant body illusion. The images for biological men were comparable, but for the figure replaced due to image rights. The images for transgender men were taken from Gonzalez (2015) with permission from the author.

were instructed to respond to images of biological men and positive attributes with the same keystroke ("I"). In incongruent blocks, transgender men were assigned the same key as positive attributes instead. Each block consisted of 16 trials.

D scores for each participant were calculated in line with Greenwald's improved scoring algorithm (Greenwald et al., 2003), which incorporates data from the practice trials, uses a metric that is calibrated by each respondent's latency variability, and includes a latency penalty for errors. Positive values represent a more positive association with biological men, and negative values with transgender.

Embodiment

For the two museum groups, directly after stimulation, participants were presented an embodiment questionnaire on a PC that was modified after previous studies assessing embodiment in similar experimental settings (e.g., Lenggenhager et al., 2007; Slater et al., 2010; Gonzalez-Franco and Peck, 2018). Dobricki and de la Rosa (2013) proposed question categories after a principal component analysis, and Gonzalez-Franco and Peck (2018) after a qualitative analysis; thus, we considered the inclusion of a question related to relevant categories to suffice for this study, given the limited time we could ask of participants attending the exhibition. The participants answered with a mouse on a VAS ranging from 0 (strongly disagree) to 100 (strongly agree) to six statements, of which three referred to illusory embodiment ("Sometimes it felt like the seen body is my own body," "Sometimes it felt like I was in control of the virtual body," "Sometimes I felt connected with the seen body," respectively), one to the sense of presence ("Sometimes I felt like I was actually there in the presented environment"), one to enjoyment of the illusion ("I found the experience enjoyable"), and one to the perceived multisensory synchrony of the event ("The seen events were synchronous to the felt events"; low scores were expected in the *conventional* group as visual and somatosensory–motor stimulation would not correspond except for head movements).

The *control* group also completed an embodiment questionnaire within the unrelated experience; however, these data are not analyzed as they are not relevant to the research questions.

Data Treatment and Statistical Analysis

Statistical analyses were conducted using JASP (version 0.11.1); the data were tested for normality, and corresponding tests were performed. An aligned ranks transformation analysis of

variance (ANOVA) for nonparametric factorial analyses (ART–ANOVA; Wobbrock et al., 2011) and the corresponding *post hoc* comparisons were performed using R version 3.5.1. Two-tailed comparisons are reported in all cases.

RESULTS

Effect of Multisensory Stimulation on Embodiment in the Immersive Experience

In **Table 1**, the questionnaire scores and Mann–Whitney independent-samples *t* test statistics are depicted. Unlike predicted in hypothesis 1, no significant differences of embodiment or enjoyment between the museum groups were found; however, a trend was found for embodiment. A marginal difference of presence and a highly significant difference of perceived synchrony were found, with higher scores in the *sensorimotor* as compared to the *conventional* group. As embodiment and sensorimotor sharing have previously been shown to depend on age (Tajadura-Jiménez et al., 2012), and embodiment did not significantly differ between the two groups, a Spearman correlation was calculated for the whole museum sample. As predicted in hypothesis 4, age was negatively correlated with embodiment ($\rho = -0.27$, $p = 0.023$; **Figure 2**), suggesting that embodiment was stronger in younger participants.

Effects of Embodying Jonah on Bias Toward Transgender People

Further explorative analyses can be found in the **Supplementary Material**.

Implicit bias: An ANOVA with the within-factor group (*sensorimotor*, *conventional*, and *control*) and between-factor time (pre, post) was performed to assess the effect of the intervention on the BIAT scores (pre: mean = -0.034, SD = 0.52; post: mean = -0.004, SD = 0.54). No significant effects were found for time [$F(1,111) = 0.47$, $p = 0.5$, $\eta = 0$; hypothesis 2], group [$F(2,111) = 0.04$, $p = 0.96$, $\eta = 0$], or the interaction between these [$F(2,111) = 0.18$, $p = 0.83$, $\eta = 0$; hypothesis 3].

Explicit bias: For the questionnaire (pre: median = 1.25, IQR = 2.56; post: median = 0.9, IQR = 2.66; these scores are given to describe the nonparametric data), an ART-ANOVA with the within-factor group and between-factor time showed no main effect of time [$F(1,111) = 3.5$, $p = 0.65$, hypothesis 2], group [$F(2,111) = 0.9$, $p = 0.39$], or an interaction between these [$F(2,111) = 1.2$, $p = 0.3$, hypothesis 3].

TABLE 1 | Median and IQR (given that data were not parametric) of the embodiment scores, Mann–Whitney comparisons between groups and rank biserial correlation (*rB)* as a measure of effect size are shown.

Variable	Sensorimotor group	Conventional group	*W*	*p*	*rB*
Embodiment	6.57 (2.07)	5.9 (2.54)	693	0.057	0.285
Enjoyment	6.8 (3.7)	5.75 (5.85)	673	0.097	0.249
Presence	6 (3.2)	5 (3.97)	698	0.049	0.294
Synchrony	7.3 (3.9)	3.6 (3)	828	<0.001	0.536

DISCUSSION

Multisensory Stimulation and Embodiment

In two groups pertaining to a sample taken during a museum exhibition, this study aimed to assess how embodying an outgroup member during either an enhanced *sensorimotor* VR experience involving movement and touch (*sensorimotor* group) or a more conventional audiovisual one (*conventional* group) would impact the perception of outgroup members. While a significantly higher embodiment score for the *sensorimotor* compared to the *conventional* group was expected (hypothesis 1), only a statistical trend in the predicted direction (together with a mild difference in presence) was found. These findings might be due to the reduced number of items in the questionnaire applied compared to other studies. However, while most studies suggest that congruent sensorimotor signals over fake bodies or limbs are fundamental to produce illusory embodiment (Botvinick and Cohen, 1998; Lenggenhager et al., 2007), it has been suggested that the induced substitution of optic flow related to head movements may be important for embodiment and even overwrite the effect of asynchronous signals in peripheral limbs (Roel Lesur et al., 2018). Here, both conditions had a congruent visual field, which might explain why only a trend was found in the data. Furthermore, while both the experimenter and participant attempted to synchronize their movements and the perceived synchrony scores were reasonably high, there might have been slight mismatches in the synchronization in the *sensorimotor* group, which may have affected the embodiment scores.

Age and Embodiment

Interestingly, in line with hypothesis 4 and previous studies (Tajadura-Jiménez et al., 2012; Graham et al., 2015), across the full museum sample (including the *sensorimotor* and *conventional* groups), a negative correlation between age and embodiment was found, with illusory embodiment decreasing with age. Bodily self-plasticity is often measured as susceptibility to illusory embodiment; these results suggest that such plasticity decreases with age. However, other studies involving the rubber hand illusion suggest no differences between younger and older adults (Campos et al., 2018; Palomo et al., 2018). It could be argued that this decrease in embodiment with age is due to the different body appearance between elder participants and the young seen body; however, there is increasing evidence that it is possible to embody radically different bodies in terms of age (Tajadura-Jiménez et al., 2017) and other features (Hoort et al., 2011). It seems that investigations of bodily self-plasticity across different adult age groups have not generally used VR or full-body illusions. Given that the body goes through significant changes in late adulthood, bodily self-plasticity as an adaptive capacity is of relevance for fundamental and applied science. Thus, extending experimental settings to contexts other than university laboratories may be

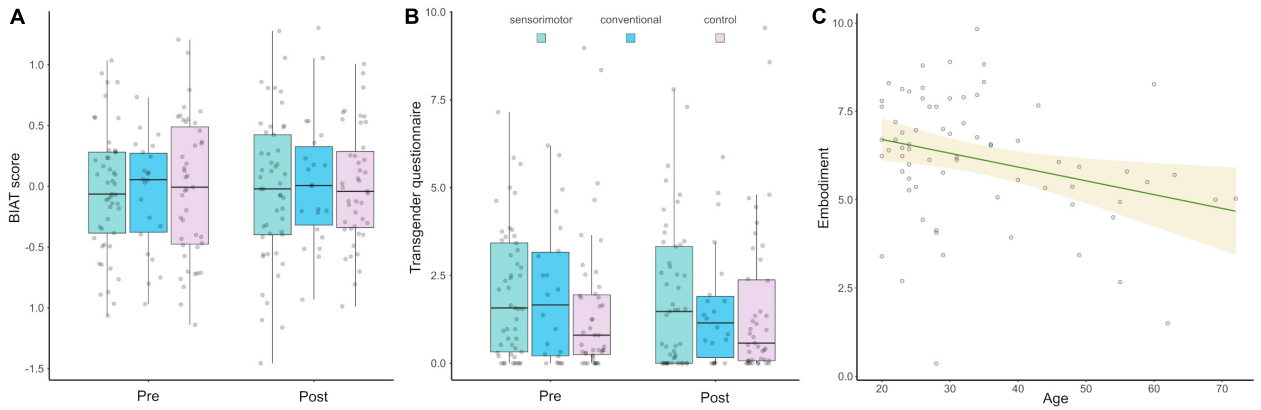

FIGURE 2 | Medians, IQR, and participants' mean scores for **(A)** the implicit measure and **(B)** the explicit measure, before and after the intervention for the three tested groups; and **(C)** negative correlation between participants' age and embodiment score with linear regression and confidence intervals for the museum sample only (which includes both the *sensorimotor* and *conventional* groups).

important, and VR might be an interesting tool to do so (Mottelson and Hornbæk, 2017).

The Effect of Illusory Embodiment on Social Cognition

Unlike expected in hypotheses 2 and 3, the reduction of bias after as compared to before the transgender experience was not found in either the implicit or explicit measure for any of the experimental groups. This finding contrasts several studies showing differences in bias toward outgroup members after embodying an outgroup member (e.g., Peck et al., 2013; Maister et al., 2015; Banakou et al., 2016). Notably, in contrast to these, the current study administered a pre–post BIAT within the same session, and the results from an explorative analysis (see **Supplementary Material**) suggest that there may have been a practice effect. While similar designs have previously presented significant results (e.g., Galli et al., 2015), practice effects could have potentially masked the intervention effect in this study. Moreover, implicit association tests are not without pitfalls (Blair et al., 2001; Fiedler et al., 2006). An alternative potential explanation for the null findings could be that while previous studies based their stimulation on a visually salient skin color, here the identity of the embodied body was not revealed visually but through the story. Future studies should look at this important conceptual difference and how it affects social cognition. Importantly, other studies found no change of implicit bias after sensorimotor sharing with an outgroup member in a virtual setting (Hasler et al., 2014). The authors suggest that this may be related to the social setting presented in the virtual environment. Others found an increase of bias after embodying a dark-skinned avatar in a social setting (Groom et al., 2009), yet these results could also be related to a population bias regardless of the experimental manipulation. Future research linking narrative and social situations to explicit bodily conditions is encouraged to disentangle the potential mechanisms that may explain the contradicting results in the literature. In this study, a floor effect is plausible (see **Supplementary Material**) because of the generally low bias in the studied population previous to the intervention, which might have diminished the effects. Unsurprisingly, the population attending the exhibition and willing to take part in a study where they embody a transgender man had low explicit and implicit bias and might not be representative of the general population.

Extending Science Through Alternative Spaces and Interdisciplinary Links

Following the above, performing this experiment in a museum setting does not overcome some of the problems intrinsic to doing research within universities (e.g., Henrich et al., 2010). Yet, the potential of VR to immerse participants in ecologically valid settings involving social situations may be still be important to further knowledge in several fields (Mottelson and Hornbæk, 2017). While methods that improve reproducibility are still needed, these types of collaborations between different disciplines and in diverse contexts may be important to move away from the artificial setting of laboratory studies. Although VR allows for a certain degree of reproducibility of the experience, the complexity of the social setting may always have an effect on participants' expectations and behavior. A museum setting, as many other ecologically valid settings, is a socially active and complex environment with many parameters out of the researcher's control. There is no clear escape from the paradox of maximally controlled settings versus ecological validity, yet VR may offer an important middle ground for future psychological research potentially benefiting from interdisciplinary collaborations. While future explorations in naturalistic settings are encouraged, they could benefit from finding public events attracting a more heterogeneous population, given that a museum might be similarly biased to a university setting.

Integrating artistic practices in scientific research might be beneficial to potentially expand scientific methods, as well as to validate certain aspects of the art methodologies

(Borgdorff et al., 2020). However, as is often the case, in order to produce reproducible science, such methodologies might be subject to simplification and homogenization. Art, in contrast to quantitative science, tends to rely on qualitative approaches assuming the variability of participants' interpretations and focus on the process rather than the outcome, potentially hindering scientific generalizations. Thus, in the transfer of artistic methods to scientific inquiry, important elements might be impaired without necessarily diminishing the legitimacy of artistic practice. For some artists, attempts aiming at the scientific legitimation of their work could imply an unwanted hierarchy between fields. Regardless of the focus of this work on experimental psychology, given its interdisciplinary nature, it seemed important to stress that potential benefits of art processes may be overlooked by the (however reproducible and robust) sometimes narrow lens of science.

DATA AVAILABILITY STATEMENT

The datasets generated for this study are available on request to the corresponding author.

ETHICS STATEMENT

The studies involving human participants were reviewed and approved by Ethics Committee of the Faculty of Arts and Social Sciences at the University of Zurich (Approval Number: 17.12.10). The patients/participants provided their written informed consent to participate in this study. 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

BL and MR designed the experiment and wrote the overall content of the article. SL contributed with the data collection. BL, MR, and SL contributed to the statistical analysis and the results section. All authors contributed to the article and approved the submitted version.

FUNDING

MR and BL were supported by the Swiss National Science Foundation (Grant Number: PP00P1_170511).

ACKNOWLEDGMENTS

We want to thank the photographer Claudia Gonzalez for allowing us to use her photo series *Reassign* for our BIAT, BeAnotherLab for creating the artistic methodology, Jonah for sharing his story and Philippe Bertrand for his contribution to the experimental design and valuable feedback for the manuscript. We also want to extend our gratitude to Jasmine Ho for proofreading our work and both Irene Madoka Scheiwiller and Marcia Arbenz for their help with data collection. Lastly, we want to thank Katharina Weikl and the 100 Ways of Thinking team for their support and invitation to participate in the exhibition.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fpsyg.2020.01861/full#supplementary-material

REFERENCES

Avenanti, A., Sirigu, A., and Aglioti, S. M. (2010). Racial bias reduces empathic sensorimotor resonance with other-race pain. *Curr. Biol.* 20, 1018–1022. doi: 10.1016/j.cub.2010.03.071

Banakou, D., Hanumanthu, P. D., and Slater, M. (2016). Virtual embodiment of white people in a black virtual body leads to a sustained reduction in their implicit racial bias. *Front. Hum. Neurosci.* 10:601. doi: 10.3389/fnhum.2016.00601

BeAnotherLab (2020). *Library of Ourselves*. Available online at: http://beanotherlab.org/home/work/library-of-ourselves/ (accessed February 11, 2020).

Bernhardt, B. C., and Singer, T. (2012). The neural basis of empathy. *Annu. Rev. Neurosci.* 35, 1–23. doi: 10.1146/annurev-neuro-062111-150536

Bertrand, P., González Franco, D., Pointeau, A., and Cherene, C. (2014). "The machine to be another—embodied telepresence using human performers," in *Proceeding of the TEI Conference 2014*, Munich.

Bertrand, P., Guegan, J., Robieux, L., McCall, C. A., and Zenasni, F. (2018). Learning empathy through virtual reality: multiple strategies for training empathy-related abilities using body ownership illusions in embodied virtual reality. *Front. Robot. AI* 5:26. doi: 10.3389/frobt.2018.00026

Billard, T. J. (2018). Attitudes toward transgender men and women: development and validation of a new measure. *Front. Psychol.* 9:387. doi: 10.3389/fpsyg.2018.00387

Blair, I. V., Ma, J. E., and Lenton, A. P. (2001). Imagining stereotypes away: the moderation of implicit stereotypes through mental imagery. *J. Pers. Soc. Psychol.* 81, 828–841. doi: 10.1037/0022-3514.81.5.828

Borgdorff, H., Peters, P., and Pinch, T. (eds) (2020). *Dialogues Between Artistic Research and Science and Technology Studies*. London: Routledge.

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

Campos, J. L., El-Khechen Richandi, G., Taati, B., and Keshavarz, B. (2018). The rubber hand illusion in healthy younger and older adults. *Multisens. Res.* 31, 537–555. doi: 10.1163/22134808-00002614

Damasio, A. R. (1994). *Descartes' Error*. New York, NY: HarperCollins.

de Waal, F. B. M., and Preston, S. D. (2017). Mammalian empathy: behavioural manifestations and neural basis. *Nat. Rev. Neurosci.* 18, 498–509. doi: 10.1038/nrn.2017.72

Decety, J. (2010). The neurodevelopment of empathy in humans. *Dev. Neurosci.* 32, 257–267. doi: 10.1159/000317771

Dijkerman, C., and Lenggenhager, B. (2018). The body and cognition: the relation between body representations and higher level cognitive and social processes. *Cortex* 104, 133–139. doi: 10.1016/j.cortex.2018.06.001

Dobricki, M., and de la Rosa, S. (2013). The structure of conscious bodily self-perception during full-body illusions. *PLoS One* 8:e83840. doi: 10.1371/journal.pone.0083840

Ehrsson, H. H. (2012). "The concept of body ownership and its relation to multisensory integration," in *The New Handbook of Multisensory Processes*, eds B. E. Stein and B. E. Stein (Cambridge, MA: MIT Press), 775–792.

Fiedler, K., Messner, C., and Bluemke, M. (2006). Unresolved problems with the "I", the "A", and the "T": a logical and psychometric critique of the Implicit Association Test (IAT). *Eur. Rev. Soc. Psychol.* 17, 74–147. doi: 10.1080/10463280600681248

Gallagher, S. (2000). Philosophical conceptions of the self: implications for cognitive science. *Trends Cogn. Sci.* 4, 14–21. doi: 10.1016/S1364-6613(99)01417-5

Gallese, V., and Sinigaglia, C. (2011). What is so special about embodied simulation? *Trends Cogn. Sci.* 15, 512–519. doi: 10.1016/j.tics.2011.09.003

Galli, G., Lenggenhager, B., Scivoletto, G., Molinari, M., and Pazzaglia, M. (2015). Don't look at my wheelchair! The plasticity of longlasting prejudice. *Med. Educ.* 49, 1239–1247. doi: 10.1111/medu.12834

Gonzalez, C. (2015). *REASSIGN*. Available online at: https://claudiagonzalez.format.com/personalwork (accessed April 21, 2020).

Gonzalez-Franco, M., and Peck, T. C. (2018). Avatar embodiment. towards a standardized questionnaire. *Front. Robot. AI* 5:74. doi: 10.3389/frobt.2018.00074

Graham, K. T., Martin-Iverson, M. T., Holmes, N. P., and Waters, F. A. (2015). The projected hand illusion: component structure in a community sample and association with demographics, cognition, and psychotic-like experiences. *Attent. Percept. Psychophys.* 77, 207–219. doi: 10.3758/s13414-014-0748-6

Greenwald, A. G., McGhee, D. E., and Schwartz, J. L. K. (1998). Measuring individual differences in implicit cognition: the implicit association test. *J. Pers. Soc. Psychol.* 74, 1464–1480. doi: 10.1037/0022-3514.74.6.1464

Greenwald, A. G., Nosek, B. A., and Banaji, M. R. (2003). Understanding and using the implicit association test: I. An improved scoring algorithm. *J. Pers. Soc. Psychol.* 85, 197–216. doi: 10.1037/0022-3514.85.2.197

Groom, V., Bailenson, J. N., and Nass, C. (2009). The influence of racial embodiment on racial bias in immersive virtual environments. *Soc. Influ.* 4, 231–248. doi: 10.1080/15534510802643750

Hasler, B. S., Hirschberger, G., Shani-Sherman, T., and Friedman, D. A. (2014). Virtual peacemakers: mimicry increases empathy in simulated contact with virtual outgroup members., Virtual peacemakers: mimicry increases empathy in simulated contact with virtual outgroup members. *Cyberpsychol. Behav. Soc. Netw.* 766, 766–771. doi: 10.1089/cyber.2014.0213

Henrich, J., Heine, S. J., and Norenzayan, A. (2010). The weirdest people in the world? *Behav. Brain Sci.* 33, 61–83. doi: 10.1017/S0140525X0999152X

Hoort, B., van der, 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

Kalckert, A., and Ehrsson, H. H. (2012). Moving a rubber hand that feels like your own: a dissociation of ownership and agency. *Front. Hum. Neurosci.* 6:40. doi: 10.3389/fnhum.2012.00040

KCTS9 (2018). *Stepping into the Body of a Trans Man*. Available online at: https://www.youtube.com/watch?v=JFW2IfuNO-Q&feature=emb_logo (accessed April 16, 2020).

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

Lenggenhager, B., Tadi, T., Metzinger, T., and Blanke, O. (2007). Video ergo sum: manipulating bodily self-consciousness. *Science* 317, 1096–1099. doi: 10.1126/science.1143439

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

Maister, L., Slater, M., Sanchez-Vives, M. V., and Tsakiris, M. (2015). Changing bodies changes minds: owning another body affects social cognition. *Trends Cogn. Sci.* 19, 6–12. doi: 10.1016/j.tics.2014.11.001

Maselli, A., and Slater, M. (2013). The building blocks of the full body ownership illusion. *Front. Hum. Neurosci.* 7:83. doi: 10.3389/fnhum.2013.00083

Mottelson, A., and Hornbæk, K. (2017). "Virtual reality studies outside the laboratory," in *Proceedings of the 23rd ACM Symposium on Virtual Reality Software and Technology*, Gothenburg, 1–10.

Palomo, P., Borrego, A., Cebolla, A., Llorens, R., Demarzo, M., and Baños, R. M. (2018). Subjective, behavioral, and physiological responses to the rubber hand illusion do not vary with age in the adult phase. *Conscious. Cogn.* 58, 90–96. doi: 10.1016/j.concog.2017.10.014

Pan, X., and Hamilton, A. F. C. (2018). Why and how to use virtual reality to study human social interaction: the challenges of exploring a new research landscape. *Br. J. Psychol.* 109, 395–417. doi: 10.1111/bjop.12290

Peck, T. C., Seinfeld, S., Aglioti, S. M., and Slater, M. (2013). Putting yourself in the skin of a black avatar reduces implicit racial bias. *Conscious. Cogn.* 22, 779–787. doi: 10.1016/j.concog.2013.04.016

Petkova, V. I., and Ehrsson, H. H. (2008). If I were you: perceptual illusion of body swapping. *PLoS One* 3:e3832. doi: 10.1371/journal.pone.0003832

Roel Lesur, M., Gaebler, M., Bertrand, P., and Lenggenhager, B. (2018). The plasticity of the bodily self: head movements in bodily illusions and their relation to gallagher's body image and body schema. *Construct. Found.* 14, 94–105.

Roel Lesur, M., Weijs, M. L., Simon, C., Kannape, O. A., and Lenggenhager, B. (2020). Psychometrics of disembodiment and its differential modulation by visuomotor and visuotactile mismatches. *IScience* 23:100901. doi: 10.1016/j.isci.2020.100901

Shamay-Tsoory, S. G., and Mendelsohn, A. (2019). Real-life neuroscience: an ecological approach to brain and behavior research. *Perspect. Psychol. Sci.* 14, 841–859. doi: 10.1177/1745691619856350

Slater, M., Spanlang, B., Sanchez-Vives, M. V., and Blanke, O. (2010). First person experience of body transfer in virtual reality. *PLoS One* 5:e10564. doi: 10.1371/journal.pone.0010564

Sriram, N., and Greenwald, A. G. (2009). The brief implicit association test. *Exp. Psychol.* 56, 283–294. doi: 10.1027/1618-3169.56.4.283

Sutherland, A. (2015). *The Limits of Virtual Reality: Debugging the Empathy Machine*. Cambridge, MA: MIT.

Tajadura-Jiménez, A., Banakou, D., Bianchi-Berthouze, N., and Slater, M. (2017). Embodiment in a child-like talking virtual body influences object size perception, self-identification, and subsequent real speaking. *Sci. Rep.* 7:9637. doi: 10.1038/s41598-017-09497-3

Tajadura-Jiménez, A., Longo, M. R., Coleman, R., and Tsakiris, M. (2012). The person in the mirror: using the enfacement illusion to investigate the experiential structure of self-identification. *Conscious. Cogn.* 21, 1725–1738. doi: 10.1016/j.concog.2012.10.004

van Baaren, R., Janssen, L., Chartrand, T. L., Dijksterhuis, and Ap. (2009). Where is the love? The social aspects of mimicry. *Philos. Trans. R. Soc. B Biol. Sci.* 364, 2381–2389. doi: 10.1098/rstb.2009.0057

Wobbrock, J. O., Findlater, L., Gergle, D., and Higgins, J. J. (2011). "The aligned rank transform for nonparametric factorial analyses using only anova procedures," in *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* Vancouver, BC, 143–146.

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: 29 October 2020
doi: 10.3389/fpsyg.2020.539726

Losing Ourselves: Active Inference, Depersonalization, and Meditation

George Deane[1], *Mark Miller*[2]* *and Sam Wilkinson*[3]

[1] Department of Philosophy, University of Edinburgh, Edinburgh, United Kingdom, [2] Department of Informatics, University of Sussex, Brighton, United Kingdom, [3] Department of Sociology, Philosophy and Anthropology, University of Exeter, Exeter, United Kingdom

Disruptions in the ordinary sense of selfhood underpin both pathological and "enlightened" states of consciousness. People suffering from depersonalization can experience the loss of a sense of self as devastating, often accompanied by intense feelings of alienation, fear, and hopelessness. However, for meditative contemplatives from various traditions, "selfless" experiences are highly sought after, being associated with enduring peace and joy. Little is understood about how these contrasting dysphoric and euphoric experiences should be conceptualized. In this paper, we propose a unified account of these selfless experiences within the active inference framework. Building on our recent active inference research, we propose an account of the experiences of selfhood as emerging from a temporally deep generative model. We go on to develop a view of the self as playing a central role in structuring ordinary experience by "tuning" agents to the counterfactually rich possibilities for action. Finally, we explore how depersonalization may result from an inferred loss of allostatic control and contrast this phenomenology with selfless experiences reported by meditation practitioners. We will show how, by beginning with a conception of self-modeling within an active inference framework, we have available to us a new way of conceptualizing the striking experiential similarities and important differences between these selfless experiences within a unifying theoretical framework. We will explore the implications for understanding and treating dissociative disorders, as well as elucidate both the therapeutic potential, and possible dangers, of meditation.

Keywords: active inference, depersonalization, meditation, error dynamics, self

OPEN ACCESS

Edited by:
Anna Ciaunica,
University of Porto, Portugal

Reviewed by:
Yasmina Jraissati,
Ronin Institute, United States
Jane Elizabeth Aspell,
Anglia Ruskin University,
United Kingdom

***Correspondence:**
Mark Miller
m.d.miller@sussex.ac.uk

Specialty section:
This article was submitted to Theoretical and Philosophical Psychology, a section of the journal Frontiers in Psychology

Received: *06 March 2020*
Accepted: *02 October 2020*
Published: *29 October 2020*

Citation:
Deane G, Miller M and Wilkinson S (2020) Losing Ourselves: Active Inference, Depersonalization, and Meditation. Front. Psychol. 11:539726. doi: 10.3389/fpsyg.2020.539726

INTRODUCTION

In daily life, we take for granted the existence of a self: we feel that we are possessors of certain qualities, the experiencers of certain sensations, that we are different and distinct from one another, and that we endure from day to day. And yet, these assumptions have long been the focus of skepticism within both Western and Eastern philosophical traditions. Thinkers from various disciplines (e.g., from philosophy of mind, cognitive science, phenomenology, and Buddhist philosophy) are beginning to collaborate on various topics revolving around self and subjectivity. One lens through which philosophers and cognitive scientists have been recently exploring the self is through cases where subjects report a loss, or diminishment, of their sense of self. These reports occur most prominently in the context of psychiatric disorders such as depersonalization (e.g., Colombetti and Ratcliffe, 2012; Seth et al., 2012; Miller et al., 2020), meditation (e.g., Britton, 2019; Lutz et al., 2019), and psychedelic drugs (e.g., Millière, 2017; Deane, 2020). The active

inference framework—a popular approach to modeling action and perception that uses principles of variational Bayesian inference (Friston et al., 2017)—is particularly promising for understanding these phenomena.

Our aim in this paper is to provide an updated account of selfless experience within the active inference framework[1]. By selfless experience, here, we mean the diminished sense of self that is reported in a wide variety of cases including depersonalization and meditative insight. Active inference and predictive processing have already been used to provide accounts of depersonalization in psychiatric contexts (Seth et al., 2012; Gerrans, 2019), and although we find these accounts promising, we seek to build on them in important ways. In particular, we differ from existing accounts in taking affective valence and control to be central to the sense of self. Building on existing accounts of self-modeling within an active inference framework (Seth, 2014; Hohwy and Michael, 2017; Friston, 2018), our account casts the self-model in terms of an allostatic control model (ACM; Deane, 2020), which we unpack in terms of "agentive control" and "motivational" components. The central thesis of this view is that the self is understood as an inference about endogenous causes of self-evidencing outcomes. In simple terms, this could be understood as the system modeling what it *wants* (motivations) and what it *can do* (abilities). However, we do not simply adopt the ACM for its own sake—there are concrete explanatory payoffs. In particular, we are better able to account for the wide range of selfless experiences under a single unifying framework. Selfless experiences come in a variety of flavors, ranging from the dysphoric and dysfunctional experiences associated with depersonalization to the euphoric and potentially superfunctional states sought after by meditators. Our explanation of this difference is, as we will see, an intrinsic part of our account of the emergence of these phenomena themselves.

An important addition to this literature that we will make is a reinterpretation of the role that affect plays in these processes. We will argue that the sense of self arises from the system's evaluation of its own performance, or predictive control, of its own adaptive behaviors. As we will see, the tracking of our performance, and the allocation of resources (i.e., setting of precision), is being done in part by affective systems. That is, we quite literally feel how well adapted we are to a situation, and those feelings move us in ways that are intended to improve that fit. This has the consequence that our sense of being a self and affect are mechanistically intertwined. This updated theoretical account of selfhood then allows us to propose a more unified framework for understanding various alterations in selfhood and affectivity.

We proceed as follows: In *From the Free Energy Principle to Hierarchical Predictive Processing*, we give an overview of the free energy principle and hierarchical predictive processing. In *A Control-Theoretic Perspective*, we position these frameworks within a control theoretic perspective and show how allostasis can be formalized in terms of active inference. In *Affect in Deep Self-Models*, we build on these ideas to present an account of self-modeling in terms of allostatic control. In *Active Inference Accounts of Depersonalization* and *ACM Account of Meditative Selflessness*, we apply this model of the self to address depersonalization and selfless experiences attained through meditation, respectively. We wrap up and conclude by comparing and contrasting these two dysphoric and euphoric selfless experiences.

FROM THE FREE ENERGY PRINCIPLE TO HIERARCHICAL PREDICTIVE PROCESSING

The free energy principle (FEP; Friston, 2010) is an ambitious unifying and overarching theory of life, according to which biological systems naturally strive to minimize free energy.

The FEP starts from the observation of existence (Friston and Stephan, 2007; Friston et al., 2010) and seeks to understand how organisms maintain their existence by "tuning" to their environmental niche, where the quantity of free energy is understood as a measure of the disattunement (which is equivalent to model "uncertainty") between the agent and environment (Bruineberg and Rietveld, 2014). Crucially, in order to exist and reproduce, agents must stay within conditions that are conducive to continued existence—such as avoiding an unacceptably high body temperature. Of course, this is *phenotype specific*—the conditions that make continued existence viable vary across species. Organisms must minimize free energy, which is equivalent to maximizing the evidence of their model and so their own existence (Friston, 2010, Hohwy, 2016). Maximizing model evidence in this way is called "self-evidencing" (Hohwy, 2016).

In animals like us (and many others), it has been proposed that free energy is minimized, at least in large part, by hierarchical predictive processing in the brain and central nervous system (Friston, 2005; Clark, 2013, 2015). What the brain has to do, on such a view, is minimize prediction error (free energy) as efficiently as possible. This requires it to come up with an overall hypothesis or model about what is going on in the world. This hierarchical model generates predictions, and if it is inaccurate, it generates *prediction error* and updates predictions accordingly.

A major challenge in model selection arises because the world is a noisy and ambiguous place. Thus, there exists, at any given time, more than one model that fits the incoming sensory signal. This is where the notion of *prior probability*, often shortened simply to *prior*, comes in (and with it, the Bayesian element of the framework). This is the background probability of the model independently of the evidence. For example, (adapting an example from Pezzulo, 2014) before I hear my downstairs front window creak open, there is a background probability concerning the likelihood that I might be burgled. Whether I live in a high- or low-crime neighborhood will influence the prior probability of the "that's a burglar!" model in response to the sound of the creaking window ("the evidence"). Models are selected based on both fit with current evidence and their prior probability. This means that you can get trade-offs, for example, where a model with a relatively low fit has a sufficiently high prior probability to

[1]See Ciaunica et al. (2020) for a recent phenomenological account of depersonalization and meditative insight.

be selected. For example, in the case of the hollow mask illusion, the model with the best fit would be the (perceptually accurate) "hollow concave face" model, but the slightly lower fit "normal convex face" model has such a high prior probability that it is selected instead, giving rise to the illusion.

This captures what the brain has to do, namely, resolve ambiguity using priors (viz., in a Bayesian manner); however, it does not tell us how this is implemented physically in the brain. Put simply, the brain maximizes efficiency (minimizes free energy) by being proactive and anticipatory. In other words, the nervous system does not passively wait for inputs to come in. Rather, even at the earliest stages of sensory processing, inputs are greeted by a barrage of top–down prediction. This does not just save time; it also saves energy and bandwidth, since the parts of the incoming sensory signals that have already been accurately predicted do not need to be passed up the processing hierarchy. All that gets passed up is what is "newsworthy" (Hosoya et al., 2005), namely, prediction error. Putting this all together, the nervous system tries to minimize prediction error by coming up with successful hierarchical predictive models that are chosen in a Bayesian manner (namely, based on fit and prior).

There are two more important tweaks to this picture. The first is to do with second-order prediction dynamics, namely, how the brain deals with statistical volatility. This requires introducing the notion of *precision*. In short, the world that we live in does not just have variability but also predictable levels of variability. As a result, our nervous systems learn over time that there are contexts where environmental information is high quality (trustworthy) and other contexts where it is not. For example, in good lighting, visual information is relatively high quality, whereas in poor lighting, it is relatively low. What an optimal system will do in response to this is have a way of setting second-order precision, namely, of appropriately varying the extent to which prediction error should be taken seriously (adjusted as a function of the likelihood of prediction error being accurate or simply noise). In high-quality informational contexts, it is expected that predictions will be good, and so prediction errors will be given relatively high weight (or gain). In low-quality contexts, prediction errors will be taken less seriously. This turning up and down of the gain on prediction error signaling is most commonly called *precision weighting*, and it plays a role far beyond the second-order dynamics that we used to introduce it. It is central to attention (Hohwy, 2012), and to the bringing about of bodily movement, an issue to which we now turn.

The second tweak comes when we note that, for embodied creatures like ourselves, action is an ever-present part of our existence. The Bayesian picture just described makes it look like we are primarily in the business of updating our models to best fit inputs from the world. However, of course, there are two ways of responding to prediction error. You can, certainly, update the model to better fit the world, but you can also update the world to better fit the model. The former is known as *perceptual inference*, and the latter is known as *active inference*. It is with the latter that you get a PP account of action and basic motivation more generally. Active inference, on our view, is central to allostasis, a notion we introduce shortly.

A CONTROL-THEORETIC PERSPECTIVE

The foundations of active inference can be traced to control theory. The idea that a system maintains existence by resisting environmental disorder by acting to remain within a limited repertoire of *phenotype-congruent* states is closely related to the notion of maintaining "essential variables" (Ashby, 2013), where an internal reference point (also known as a setpoint or goal signal) is compared to the current state and the system acts so as to restore conditions to the setpoint.

The principles of *control-oriented predictive regulation* (Seth and Tsakiris, 2018) are very similar.

Here, the brain applies the same inferential machinery of hierarchical predictive processing to infer and track key homeostatic variables, using prior expectations and afferent sensory information about the body coming "from within" (Craig, 2003). In order to stay alive, organisms have to execute the right actions to bring about state transitions that bring bodily states into reasonable bounds (Pezzulo et al., 2015). The phylogenetically endowed high precision on expectations for staying within homeostatically viable states means that the organism acts to realize prior beliefs corresponding to the maintenance of essential variables ("goal priors"), for example, eating to restore a blood sugar concentration to expected levels. While goal priors originate in the maintenance of essential variables (e.g., steady temperature, blood sugar levels, etc.), over the course of ontogeny, an organism can acquire new goal priors that are predictive on longer timescales of being relevant for maintaining homeostasis—such as staying within a particular social milieu (Matthews and Tye, 2019).

Active inference, then, formalizes homeostasis through a control theoretic lens. Homeostasis from this perspective is maintained not only through autonomic reflexes (i.e., sweating to cool down) but also by *prospective control*. Such systems anticipate future dyshomeostatic conditions before they arise and proactively act to avoid them. This prospective control relates to both inferences about current and future bodily states contingent on certain actions (Sterling, 2012; Seth, 2014; Pezzulo et al., 2015). This process of anticipatory action, by which the brain regulates the needs of the body, is known as *allostasis* (Corcoran et al., 2019). Active inference formally articulates allostasis, such that agents *anticipate* surprising outcomes before they arise and act in order to minimize uncertainty about potential future outcomes (Sterling, 2012; Pezzulo et al., 2015, 2018).

On the active inference formulation, the action selection process itself is cast as a problem of inference, where agents must infer the active sampling of the world that realizes prior preferences and minimizes uncertainty (Kaplan and Friston, 2018). Action selection, then, depends on the use of a deep temporal model, where policies (sequences of actions) are selected based on prior expectations of the quantity of free energy that the agent expects itself to average over time ("expected free energy") *given* a particular policy or course of action (Pezzulo et al., 2015; Friston et al., 2017). Intuitively, some courses of action (such as riding in the train carriage) have lower expected free energy than others (such as riding on the roof). Crucially, this involves anticipating unfavorable

or dyshomeostatic conditions before they arise and acting to minimize uncertainty about potential future outcomes (Friston et al., 2015, 2017). On this account, higher levels of the cortical hierarchy, tracking regularities unfolding on longer timescales (Kiebel et al., 2008), contextualize lower levels by anticipating the downstream consequences of action and selecting policies that minimize expected free energy according to these expectations (Friston, 2010; Pezzulo et al., 2015). For example, my longer-term goal of successfully catching the train includes expectations about what I need to do to get to the station on time, which in turn unpacks into subgoals such as getting in my car and lower-level action–prediction loops as I use the pedal, gearstick, and so on.

In order to minimize free energy over longer timescales, active inference requires balancing the *pragmatic* and *epistemic* value of different actions. The pragmatic (or instrumental) value of an action or action policy (a sequence of actions) refers to the probability of it resulting in sensory states that fulfill some prior preference or goal state, such as maintaining a viable body temperature. Epistemic value refers to the reduction in uncertainty or information gain expected under a given action or action policy (Kaplan and Friston, 2018). Epistemic action allows organisms to increase an agent's ability to reduce free energy by increasing their understanding of the predictable aspects of the environment. Information-seeking behavior such as novelty seeking and curiosity can be accounted for within this formulation in terms of epistemic action (Friston et al., 2015; Mirza et al., 2016; Kiverstein et al., 2017; Kaplan and Friston, 2018; Pezzulo and Nolfi, 2019). Intrinsic motivation (and epistemic foraging) can be understood here in terms of uncertainty reduction (Barto, 2013). Simulations of economic decision making and epistemic foraging behavior have been built based on this view that the probability of a policy is proportional to expected free energy (Friston et al., 2014, 2015, 2017). Active inference formulations of planning and navigation have been used to dissolve the "explore–exploit" dilemma, as the agent simply needs to act so as to minimize uncertainty (i.e., free energy; Kaplan and Friston, 2018). Agents engaging active inference do not just keep themselves in the states that are expected; rather, they anticipate in order to minimize uncertainty about potential future outcomes (Schwartenbeck et al., 2013; Friston et al., 2014, 2015, 2017).

Now that we have introduced the control-theoretic notion of allostasis, and how it is achieved via active inference, we next go on to develop our view of the sense of self.

THE SENSE OF SELF AS A MODEL OF ALLOSTATIC CONTROL

This section will argue that the self is best understood in terms of an *allostatic control model* (ACM). Recently, a number of computational models of the minimal sense of self (namely, the self as implicitly present in everyday world-directed experience, rather than something more overt and explicit like the self-conception or narrative self) have been advanced in the active inference literature (Limanowski and Blankenburg, 2013; Seth, 2013; Apps and Tsakiris, 2014; Allen and Friston, 2016). Common to these proposals is that the sense of self arises inferentially within a hierarchical generative model. Our central claim is that the inferential self-model arises from the system tracking its own self-evidencing capabilities (Friston, 2018). The purpose of tracking these capacities is to infer confidence (precision) in potential action policies according to their expected free energy and thereby arbitrate between potential actions accordingly. Self-modeling of this kind, then, is fundamentally related to selecting allostatic or anticipatory actions, where the system preemptively infers and avoids unfavorable conditions before they arise. By casting the self-model in terms of allostatic control, we will connect our view in new ways to the prevalent theme in neuroscience about the rich relationship between affectivity and the self (Damasio, 2003; Seth, 2013; Allen and Tsakiris, 2018). This view can be understood formally in terms of a higher-level inference about "subjective fitness"—that is, a higher level of the generative model that scores the "fit" between the action model and the world (see Hesp et al., 2019 for a formal treatment and computational model of this idea). Conceptually, our view of the sense of self can be decomposed into "agentive control" and "motivational" components. We will present these in turn.

Agentive Control

Recall that while perception involves updating the model to better predict the incoming sensory input, action changes the incoming sensory input to better fit the model. In selecting an action, then, the system implicitly infers itself as able to bring about the consequences of that action. A sense of agency, the sense of being the one in control of an action, naturally emerges here as part of model sampling (Friston et al., 2013)—in selecting an action, the system implicitly infers itself as able to bring about the sensory consequences of the action. On this view, the sense of control—the expectation of being able to bring about certain consequences given certain actions—is *learned* through past experiences of the system inferring its own agentive capacities.

This connects closely with preexisting accounts of the sense of agency, where the system infers its own agency based on the ability to predict the outcome of a given action (Haggard, 2017). Here, attribution to endogenous causes (self), as opposed to exogenous causes (world/other), occurs as the result of a "comparator model," where the sensory consequences of an action are compared with the expected sensory consequences (Frith, 2014). This allows the system to sculpt and improve motor control, as the discrepancy between the sensory consequences of an action are compared with the predicted (intended) outcomes. The system can then act to iteratively reduce this discrepancy and refine motor commands (Miall and Wolpert, 1996; Wolpert and Flanagan, 2001).

Crucially for the current account, *control* is *temporally deep* (Pezzulo, 2018), such that the agent not only has predictions about the immediate consequences of actions but also of consequences extending into the future. The sensory consequences of a given action may be sensorially proximal (e.g., the immediate sensory consequences of hitting send on an email), or sensorially distal and abstract (e.g., the expectation that when you see that person they will know the information in the email).

The system, then, must be able to track the outcomes of actions on multiple timescales. Within the generative model, lower and higher levels of the hierarchy track regularities unfolding at faster and slower timescales, respectively (Kiebel et al., 2008). For an organism with a temporally deep generative model, this includes tracking its expected control of actions on short timescales (e.g., the expected sensory consequences of taking a step) and using these inferences to inform inferences about the state of control on temporally deep timescales (e.g., being able to walk a distance). On this view, the system models itself as an agent according to this hierarchically deep inference about its own endogenous control of sensation via its actions. In other words, I have a sense of what I can do based on past experience of acting in the world and come to expect myself as a controller over my future actions.

Motivation

The *motivational* component of our view of the self-model is understood in terms of *goal priors* and, as such, connects closely to views of selfhood grounded in interoception (Seth and Friston, 2016; Barrett, 2017; Seth and Tsakiris, 2018). This is because the system will be generally more concerned about controlling the "essential variables" (i.e., homeostatic set points like blood sugar levels) tracked by interoception than variables inferred through exteroception and proprioception, which are less likely to pertain directly to homeostasis (Seth, 2014; Seth and Tsakiris, 2018).

Creatures tracking longer timescales can augment this with *deep goal hierarchies* (see Pezzulo et al., 2018), where fulfillment of longer-term goals can be traded off with fulfillment of shorter-term goals. On this view, low-level maintenance of "essential variables" are phylogenetically endowed expectations that, due to an "*a priori* hyperprecision of visceral channels" (Allen and Friston, 2016, p. 7), the system must act to fulfill, rather than simply updating via perceptual inference. One example would be moving to the shade under a tree to maintain viable body temperature. Divergence from these fundamental, phenotype-congruent low-level prior expectations tunes attention and amplification of sensory signals. This manifests itself to the system as, for example, the feeling of hunger (interoceptive prediction error) or a violation of the "healthy body condition" prior in the case of pain (Ongaro and Kaptchuk, 2019). These interoceptive changes tune the organism to the appropriate action opportunities in the given context, such as finding food to resolve interoceptive prediction errors or removing the source of pain. Crucially, pain is tuned relative to expectations given the context (Moutoussis et al., 2014). In the case of an approaching bear, the prospective inference about imminent catastrophic prediction error of being eaten trumps the proximal pain of a twisted ankle, and the selected policy is running away. Put another way, hierarchically deep contextualization of interoceptive signals tunes an organism to appropriate actions and engagements with the environment (Pezzulo and Cisek, 2016) and assigns appropriate precision to priors and ascending prediction errors. For low-level drives and motivations, this is intuitive—the hungry organism is tuned to capitalize on eating opportunities present in the environment. Precision on goals tracking different timescales are continually being traded off between levels—such as refraining from eating chocolate cake in the present for the sake of a longer-term goal of sticking to a diet (Pezzulo et al., 2018).

The sense of self, then, emerges as the result of a hierarchically deep inference about the system's control of its own self-evidencing outcomes. In the generative model, this means that the sense of self can be understood in terms of a higher-level inference about the "fit" between the current action model and the world. By fit here we mean how well or poorly one is doing at reducing error over time relative to expectations. As we will see in the next section, a key implication of this picture is that self-modeling is fundamentally affective, where affective changes in the body tracks how well the organism is doing at fulfilling its own goal priors ("subjective fitness;" Joffily and Coricelli, 2013; Seth and Friston, 2016; Kiverstein et al., 2017). An upshot of this picture is that self-modeling, and indeed the feeling of being a self, is connected to affect in ways previously underappreciated in the literature, as we will explore next.

Affect in Deep Self-Models

The previous section argued for a view of self-modeling as a higher-level inference about the system's allostatic control. In our view, as we will now see, minimal self-modeling and affect are coconstitutive, such that affect can be understood as an inference about the performance of the action model in bringing about self-evidencing outcomes. This section unpacks how inference about allostatic control, manifesting affectively, is central to the allocation of precision.

In tracking the performance or "fitness" of the model over time, the system becomes sensitive to *the rate of error reduction.* In selecting a policy, the system has prior expectations of the rate at which error is likely to be reduced over time. The system can then evaluate whether its performance at reducing error is better or worse relative to its prior expectations. We can think of each agent's performance in reducing error then in terms of a slope that plots the various speeds that prediction errors are being accommodated relative to their expectations. Changes in the rate at which error is reduced (referred to as "error dynamics") turns out to be an important source of information for a predictive organism, as it reflects the efficiency, and so the quality, of its action model performance over time. As such, error dynamics play an important role in tuning precision estimations—increasing or decreasing our beliefs in the reliability of the model generating the policy (Kiverstein et al., 2017; Hesp et al., 2019). If precision is set based on estimations of how likely some action is to lead to the expected result, then the efficiency—the rate at which error is reduced—of those actions to reduce error should be taken into consideration. Greater than expected error for a given policy is evidence that the system should downregulate precision on the action model. Sensitivity to error dynamics increases our capacity to reduce prediction error over longer timescales, as it affords a means to toggle confidence levels on the action model according to the volatility of the environment.

The phenomenological manifestation of this (subpersonal) sensitivity to error reduction rates over time is affect. There is a growing literature that supports the view that affective changes not only track changes in immediate divergences from

the homeostatic ideal, as was the focus of earlier predictive accounts of interoception (see Seth, 2013), but also tracks the rate of change in error management over time (Joffily and Coricelli, 2013; Kiverstein et al., 2017; Van de Cruys, 2017). Valenced bodily feelings (i.e., positive and negative hedonic tone) are, in part, a reflection of how well or poorly we are reducing error over time relative to expectations. When error is being reduced slower than expected, and the organism is becoming increasingly disattuned to its environment, this change is marked by feelings of frustration and disappointment. The negatively valenced bodily feelings provide the organism with feedback about the reliability of the selected action policies, indicating a need to downregulate precision on those policies. In contrast, when error is being managed at a better than expected rate, the organism is gripping the scene well, the bodily feedback are positive feelings of hope and satisfaction, and precision is upregulated. It is intuitive that persistently worse than expected rates of error reduction on a given goal prior act as a disincentive to pursue that goal and motivate the system to select a more achievable goal, and doing well is motivating to continue to realize a certain goal. Precision does not just concern the organism here and now and its momentary state of uncertainty but is instead helping it to continuously improve working toward managing uncertainty over time. Importantly, positive and negative feelings alter precision relative to the rate at which we have come to expect errors to be resolved.

Affective valence here is being reimagined within the active inference framework as a domain general controller that tracks and assigns precision relative to changes in our expected rates of error reduction (that is, expected reductions in free energy; Kiverstein et al., 2017; Hesp et al., 2019). Inference about how well the system is self-evidencing as a whole is tracking a long-term dimension of the self, which is necessarily more invariant and abstract in virtue of tracking a longer timescale, showing less variability than "lower" aspects of the self-model that are more amenable to changing across contexts. Negative and positive feelings then track lesser than expected and greater than expected allostatic control, respectively. This higher-order inference about the system's confidence in its own action model, used to modulate precision on expected free energy (Hesp et al., 2019), is a candidate computational correlate for the sense of self—the feeling of being an agent.

This account of self-modeling as mechanistically intertwined with affectivity is, at present, a theoretical proposal. However, recent work (most notably Hesp et al., 2019) provides proof of principle of how this theoretical framework can be modeled computationally. This is a very promising groundwork for future work in computational modeling that is able to tie both phenomenology and behavior to underlying computational mechanisms. An important consequence of highlighting this underappreciated link between affect and self-modeling in active inference is that it provides a bridge between these computational frameworks and the phenomenology of being a self (for another account of this, see Kiverstein et al., 2020). Bodily feelings here represent a prereflective source of information about how well an agent is doing in their predictive engagements. These feelings give them a sense of what they can do, of what is possible, and what is not possible (the sense of "I can"). We have a feel for what is possible in the world based on what we can do in the particular situation we find ourselves within. Above, we characterized bodily feelings as driving policy selection. The result is that one quite literally feels drawn to relevant action possibilities. These bodily feelings track which possibilities are relevant to an agent and move us to improve[2]. The result is an ongoing dynamic dialectic between agent and environment all circling around affectivity.

While the importance of this ongoing tension between bodily feeling and environmental affordances is easily overlooked when it is functioning well, alterations in this quality can have devastating effects on how one experiences oneself and one's world. With the addition of these more recent computational models of valence as setting precision relative to changes in control, we have now for the first time at our disposal the means to provide the fullest expression of an active inference account of the sense of self. In the rest of this paper, we will use this more fully realized view of the self to propose a new unified account of the alterations in self-experience native to depersonalization and meditative insight.

ACTIVE INFERENCE ACCOUNTS OF DEPERSONALIZATION

Depersonalization disorder (DPD) is still a relatively neglected dissociative disorder. Dissociative disorders are a class of mental illness characterized by disruptions in perception, consciousness, and/or identity. These disruptions can cause various symptoms that are problematic for a person's life including social relationships and work life. Recently, however, research into the phenomenon of depersonalization more generally is increasing in part due to the piqued interest of philosophers and cognitive scientists interested in the nature and function of the self. Depersonalization experiences potentially provide researchers with important glimpses into the neuropsychological mechanisms and functional profiles of our ordinary experiences of being a self (see Metzinger on philosophy and dissociative disorders). A hallmark of depersonalization is a disturbance in subjective experience. This commonly includes a sense of detachment or alienation toward themselves, their bodies, and their environments. While specific disturbances in self-related experiences (depersonalization) and their experience of the environment (derealization) can come apart, they commonly co-occur; we will have more to say about his co-occurrence shortly (Sierra and David, 2011).

London-based writer Gracie Lofthouse writes on her own experience of depersonalization in a recent article:

> *"The first time I can remember feeling like I didn't exist, I was 15. I was sitting on a train and all of a sudden I felt like I'd been dropped into someone else's body. My memories, experiences, and feelings—the things that make up my intrinsic sense of "me-ness"—projected*

[2]For an excellent account of the neuroscience supporting the role of affect (including valence and arousal) in simultaneously tracking the relationship between the organism and the environment, and preparing the organism to make improvements to that relationship, see Lisa Feldman-Barrett's work (2017).

across my mind like phantasmagoria, but I felt like they belonged to someone else. Like I was experiencing life in the third person" (Lofthouse, 2014).

Most people have some experience of this sort of state. If you have not, it can be difficult to understand, and indeed, sufferers of depersonalization commonly report difficulties in expressing their experiences (Simeon and Abugel, 2006, p. 80). Depersonalization symptoms can last for moments, or several years, and commonly accompany major depression, anxiety disorders, substance addiction, brain injury and disease, and emotional trauma. An increasingly popular view of depersonalization is that it may act like an "airbag" in traumatic situations: when fight or flight are unable to remove an overwhelming, emotionally painful, experience, then the affective system may have its volume turned down as a direct means of reducing the suffering. The result of this reduction is what Medford calls it "desomaticion" or "deaffectation" (Sierra et al., 2005; Simeon et al., 2008; Medford, 2012; Medford et al., 2016) and is potentially the cause of the characteristically strange phenomenon of losing something important about the self and the world (see, e.g., Sierra and Berrios, 1998; Radovic and Radovic, 2002; Medford et al., 2005; Sierra et al., 2005; Simeon and Abugel, 2006; Baker et al., 2007).

In our view, predictive processing has offered some of the most promising avenues for understanding depersonalization. Our main aim here is to build on these and to improve on them based on more recent developments in the literature on active inference and the view of the self-model outlined in the previous section. The main explanatory payoffs that we can see are not only that we can better explain depersonalization and related symptoms but that we are also well-placed to explain why some instances of loss of self can have a positive valence, while others do not. Ultimately, superficial similarities in what are described as experiences of "loss of self" mask deep underlying differences.

Existing Predictive Processing Accounts of Depersonalization

Seth et al. (2012) were perhaps the first to apply predictive processing to depersonalization, and since then, Gerrans (2019) has also provided an account. Seth et al. (2012) build their account on the central notion of "conscious presence." Since "presence" involves both a sense of oneself as present in the world, and the world as present to us, it casts depersonalization and derealization as two sides of the same coin. To briefly summarize their account, they build on work in schizophrenia research on the loss of the sense of agency (e.g., Frith, 1987; Blakemore et al., 2000) according to which this arises from imprecise predictions about the sensory consequences of actions (see also *Agentive Control*, above). This account gets adapted to account for presence. According to Seth et al. (2012),

"presence is the result of successful suppression by top-down predictions of informative interoceptive signals evoked (directly) by autonomic control signals and (indirectly) by bodily responses to afferent sensory signals. According to the model, disorders of presence (as in DPD) follow from pathologically imprecise interoceptive predictive signals." (p. 2)

Our account builds on this in a number of respects. First, this account is based on a view of emotion as interoceptive inference. So is ours, in a sense, but what Seth and colleagues mean is emotion as interoceptive *perceptual* inference. In other words, as they explicitly state (p. 1), they are fleshing out the James–Lange theory of emotion (James, 1890) according to which emotion is perception of bodily (specifically visceral) change. Given a PP gloss, whereas perception is the result of model selection for minimizing prediction error from sense perception, emotion is simply model selection for minimizing prediction error from interoception. This is perceptual inference since the model has to accommodate the input. Building on our recent work (Miller and Clark, 2017; Wilkinson et al., 2019), we, in contrast, view emotion, and affect more generally, as involving *active* inference, too. In terms of ACM, it is a central part of allostasis. This brings us to another crucial difference with our view. The view of emotion as interoceptive model building tells us nothing about *valence*. And yet, emotion has valence: it tends to be either positive or negative (and to greater or lesser degrees). In contrast, we tie positive valence to allostatic control (and negative valence to lack of such control). This means, crucially, given what we say later, that valence falls naturally out of our account, both of affect in general but also of self-loss, both negative and positive, in particular.

Unlike Seth and colleagues' account, Gerrans' account does not take presence as a basic notion out of which both self and self-loss emerge for free. Instead, Gerrans appeals to the notion of a "self-model" or, perhaps more accurately, the idea that the self-features as part of an overall predictive model that determines conscious experience. Gerrans' main point is that our ordinary experience of the world, and ourselves, is generated by a constant integration of cognitive, perceptual, and affective signals. Building on his earlier work with Chris Letheby (Letheby and Gerrans, 2017), the self here is part of a predictive model, one that works to explain away the affective changes that occur as the organism engages with its environment. When affective signals go missing, the predictive system needs to explain the absence.

Gerrans' approach to explaining depersonalization focuses on the role of affect in the generation of our felt sense of presence (Seth, 2013). Like Seth et al. (2012), Gerrans concludes that when predictions about ordinary affective reactions are not fulfilled, the system generates the sense of the agent being no longer present in the experience. In short, Gerrans builds on Seth et al. but adds the self-more explicitly into the model. To the extent that Gerrans' account is similar to Seth and colleagues' account, it shares many of the same differences with our own. Nevertheless, we like the embellishment of adding the self as a feature of the predictive model. In a sense, we would agree with Gerrans that presence emerges from a basic notion of self rather than the other way around.

What both of these existing accounts have in common, with respect to depersonalization, is the focus on affective numbing based on alterations (viz., inaccuracies) to interoceptive predictive processing. We do not disagree. However, what we add to the picture is the idea that affect carries inbuilt valence, involves active inference (allostasis), and, crucially, plays a role in setting precision weighting. This has the welcome side effect of

allowing us to neatly explain other features of depersonalization beyond simply affective numbing. It also generates an account of depersonalization according to which it is inherently a negative experience (rather than something that needs to be appraised as such after the fact). These two other accounts tell us why there might be a loss of sense of self but not why that is negative. Given the existence of extremely positive experiences of self-loss ("enlightened" states), the negativity of the experience is not something that should be taken for granted. Our explanation of this large difference in valence is that superficial similarities are masking quite radical differences in what is going on in the two cases.

In the next section, we will propose that ACM can do a better job at accounting for depersonalization experiences than previous PP accounts. In particular, we will develop a view of depersonalization as a loss of allostatic control.

ACM Account of Depersonalization

Recall that the ACM casts the sense of self as underpinned by an inference about the system's endogenous control of self-evidencing outcomes. The highest levels of the self-model are also the most enduring, due to their being the most invariant across contexts. These higher levels that track how well we are able to control our interactions with the environment (i.e., allostatic control), more generally, act as hyper-priors informing more domain general precision estimations. The result of the sense of self being hierarchically deep in this way is that a temporary loss of control within a particular context may not necessarily reduce a more general sense of control or a sense of control across contexts. In other words, someone can fail to play the violin well without losing confidence in their ability to live a good life.

This inference about allostatic control—manifesting as affective valence—plays a role in setting precision relative to changes in how well or poorly we are doing at reducing error given the context. Sensitivity to unexpected increases in prediction error rates—manifesting here phenomenologically as negative valence—acts as a disincentive to continue operating in a particular context (Kiverstein et al., 2017; Hesp et al., 2019). For example, when learning an instrument, if a certain song is too complex given our skill level, the feelings of frustration that arise could motivate task switching perhaps to a simpler song or to developing some of the skills necessary to eventually play the more complex tune. Task switching here offers a way for the system to get back to reducing error at a better rate.

However, what happens when the system cannot resolve the negative affect through task switching? In other words, how would such a system behave if unexpected error continued to rise regardless of perceptual updates and behavioral interventions? For example, in active inference terms, trauma could be understood as a massive influx of prediction error causing the system to drastically lower confidence (precision) in its action models (see Linson et al., 2020). In the case of physical trauma, the body's integrity, which is highly expected, is seriously disrupted or damaged. The system in this situation is unable to reduce errors either by updating their models (perceptual inference) or acting in a way that will bring their expectations back in line with the current situation (active inference). This disparity between expected control of prediction error and the error-riddled reality produces huge amounts of negative affect, which, as we have discussed above, reduces certainty on the currently selected policy as a means of tuning the agent to better predictive opportunities.

If an external situation continues to create error (i.e., severe pain) over an extended period of time, and the agent cannot control the situation through switching domain or context (that would otherwise be controllable), the resulting drop in precision on expected free energy is going to be such that consequent transitions between higher level affective states will be forced into the same fearful state continuously. The ascending message from the negative "affective charge" (Hesp et al., 2019) will override the descending message from higher level policies (i.e., our ability to control error by task switching also fails to resolve the issue, and so we lose confidence domain general control). That means that you have exactly the same effect at the next level, whereby the desired state (positive affect) is never reached despite trying to control it, leading to a drop in the precision on expected free energy at that level, and so on upwards. Crucially though, for this to happen, the person would have to never give up the resistance to the fear/pain, i.e., maintain a high precision on that goal state (i.e., the phylogenetically expectation to have a healthy, well-functioning body). In time, the "hopeless" situation might eventually create a learned belief that, no matter what they do, they will always be in a negative valence state (i.e., valence state transitions are not conditioned by policies and are stable in "negative"). It would basically be a perfect storm for a gridlock situation where you have a strong preference against the negative state but then you also know that you cannot escape it no matter what you do. The only option is to dissolve the process that is creating the negative affect in the first place since nothing else can work. The consequence would be an unraveling of the process by which we form affective states (i.e., inferring confidence in expected free energy) and with it the sense of self[3].

The dampening of affect and the feeling of self-loss are intimately related here. In losing a sense of allostatic control, the system ceases to posit itself as a causally efficacious controller of sensations. Accordingly, in losing a sense that the system has allostatic control across contexts, the affective system ceases to tune to opportunities to reduce error. Nothing is motivationally salient because the system infers that it is not causally efficacious in bringing about self-evidencing outcomes, and as such, it infers a global loss of confidence in precision on action policies (see Kiverstein et al., 2020). The result would be that the world would lose some of its *phenomenal depth*—it would cease to solicit one's engagements and so be perceived (just as DPD sufferers suggest) as two-dimensional, or flat (Medford et al., 2006, p. 93). The result is, as Colombetti and Ratcliffe write, that "The world ceases to matter, people and events are not salient anymore. With this, the world ceases to move and affect one through one's body" (2012, p. 148; see also work by Sass and Parnas, 2003; Fuchs, 2005).

This proposal casts new light on various circumstances of occurrence surrounding experiences of depersonalization. For

[3]Thanks to Lars Sandved-Smith for discussions about the computational nature of depersonalization.

example, consider a traumatic stressor such as torture, which is perhaps the most reliable instigator of depersonalization (Kira et al., 2013). On the current account, sustained inefficiency of the motivational system (in this case, severe pain over an extended period of time) in a scenario where the person has no control to act to resolve the prediction error results in a global loss of confidence in "tuning" affective responses. Another example is major depression. Depression can be understood as a domain general inference of loss of allostatic control, where, in extreme cases, the system ceases to posit itself as a causally efficacious agent (Fabry, 2019; Kiverstein et al., 2020). The chronic stress of an uncertain or volatile environment means eventually that the system ceases to posit endogenous control on the outcomes as the occurrence of positive/negative outcomes is inferred to be independent of the agent's actions. Through the current framework, the comorbidity of major depression with depersonalization can be understood to be a sustained loss of allostatic self-efficacy. Stephan and colleagues proposed that "the performance of interoceptive-allostatic circuitry is monitored by a metacognitive layer that updates beliefs about the brain's capacity to successfully regulate bodily states" (Stephan et al., 2016, p. 1), which they dub "allostatic self-efficacy." Other accounts have proposed that depression functions as a means of reducing prediction error associated with adverse social contexts (Badcock et al., 2017). If this is along the right lines, depression itself would function as a means of motivating withdrawal from potentially aversive contexts. The link between major depression and depersonalization can be understood here in terms of the system ceasing to posit itself as an endogenous controller of self-evidencing outcomes—when there are no possible context to move to (no high level, temporally deep goal priors to realize), the total loss of allostatic control is experienced as depersonalization[4].

In the next section, we turn our attention to an example of a potentially euphoric selfless experience, namely, the sorts of selfless experiences that can arise from meditation. A view of the self-model in terms of allostatic control gives a fitting account of this and shows how this kind of selfless experience is radically different (indeed, relative to control, they are diametrically opposed) to selfless experience in depersonalization.

ACM ACCOUNT OF MEDITATIVE SELFLESSNESS

While meditation is something of an umbrella term, the disciplined control of attention is central to almost all styles of meditation (Albahari, 2009; Austin, 2013; Garfield, 2015; Millière et al., 2018). For brevity, we will focus here specifically on *focused attention meditation*. Meditation here takes the form of consciously attending to a particular object (i.e., bodily sensation or breathing), and when the mind wanders from the chosen target and the practitioner realizes the shift, they actively "let go" of the distractor and reorient their attention back to the initial meditative target.

Active inference has recently begun to be applied to thinking about meditation (Farb et al., 2015; Pagnoni and Guareschi, 2015). Lutz et al. (2019) have provided the first account of *focused attention* meditation in terms of active inference. Focused attention meditation is described as having two interrelated aims: the pragmatic activity of regulating attention on a particular object (i.e., the breath sensations, an object) and the epistemological activity of increasing one's understanding of the nature of the meditative object and the various distractors, in particular recognizing their dynamic and impersonal nature (Lutz, 28).

In active inference terms, the pragmatic aspect of focused attention meditation requires that top–down directed precision enhance the behavioral policies associated with stable attention on an object. The challenge for the meditator then is to maintain this policy, although multiple other policies may be simultaneously active and acting as competitors for selection (Pezzulo and Cisek, 2016). In meditation, this ongoing competition becomes simplified to include only the policy of maintaining attention on the meditation object and all other competing policies attempting to divert attentional resources elsewhere (e.g., spontaneous memories, future planning, homeostatic concerns, mind wandering, etc.). Inaction during this process is considered crucial, as it is the process of setting top–down precision on the sensory signals associated with the meditative object that allows this dialectic between focused attention and distraction to unfold and to be consciously attended to. Lutz and colleagues suggest that this quality of "inaction" corresponds to the subjective experience of "letting go" of the various distractions (p. 28).

Over time, the meditator can learn to allow the various distracting thoughts and sensations to arise and pass without disturbing their concentration, that is, without disrupting the meditation policy of focused attention. In part, this occurs through learning to actively reduce precision on the distracting (negative) goal prior. Inaction itself does this to some degree. An itch motivates, via negative valence, a scratching policy because of the preference for the non-itching state (i.e., goal prior) and the high precision on the itching state (i.e., resistance to the sensations), as well as the learned connection between the itch and the action policy scratching. By not acting, and calmly observing the itch, the negative preference loses precision (i.e., resistance to the sensations is dropped). The result is that the probability of the sensation driving the selection of a new (distracted) policy is also lessened. In other words, meditators can actively reduce certainty on goal states and mitigate involuntary policy selection through inaction. The result is that over time, such goal priors cease to draw processing (i.e., attention) in the same way.

In line with our view that the system infers its own control in terms of correspondence of action–outcome contingencies, here, the system learns endogenous control on the precision of goal priors through repeated reduction via opting for the focused attention policy. Over time, the decrease in distractibility can be understood as an increased ability to endogenously

[4]Interestingly, this account is suggestive that the hyperreflective tendency to check one's body and one's current state common in people suffering from depersonalization (Colombetti and Ratcliffe, 2012) could be understood as compensatory behavior aimed at reducing the uncertainty relating to the perceived loss of control over interoceptive states.

control precision on goal priors. The system here refines attentional selection through mental action in a way analogous to refining motor commands through iterative inference of control of action–outcome contingencies (Miall and Wolpert, 1996; Wolpert and Flanagan, 2001). Note that learning to actively adjust precision on goal priors in this way is to learn to exert exactly the kind of control that is missing or weakened in cases of depersonalization as characterized above. In those cases, an inability to reduce precision on a goal prior that was unattainable leads to tremendous suffering and a loss of confidence in our ability to control the world more generally. In learning that it can endogenously set precision, it now begins to infer a domain general sense of being able to realize its goal states. This increase in domain general control may correspond to the pervasive feelings of joy and peace that characterize long-term meditation (Dambrun and Ricard, 2011; Dambrun, 2016).

During the course of becoming an adept meditator, one develops the ability to remain poised between the object of attention and the ongoing flow of spontaneous mental activities. This capacity to remain subtly focused on the meditative object, while at the same time observing clearly the spontaneous mental activity, provides an optimal opportunity to learn about the process of policy selection taking place within its own system (Ridderinkhof et al., 2004). This is the epistemological element of focused meditation. The result of this introspective investigation is a gradual "opacification" of these mental processes (Carter et al., 2005). A mental process is considered transparent insofar as its contents are available to consciousness, while its non-intentional structure or construction process are not (Metzinger, 2003). Without having access to the earlier stages of processing, transparent processes are presented subjectively as fundamentally real and personally essential. Metzinger writes,

> *"Transparent phenomenal states make their representational content appear as irrevocably real, as something the existence of which you cannot doubt. Put more precisely, you may certainly be able cognitively to have doubts about its existence, but according to subjective experience this phenomenal content—the awfulness of pain, the fact that it is your own pain—is not something you can distance yourself from. The phenomenol-ogy of transparency is the phenomenology of direct realism and in the domain of self-representation it creates the phenomenology of identification." (2017, p. 248).*

By continually letting go, the distracting policy selections are observed as non-essential—they arise, they persist for some time, and they eventually dissolve without the need for overt actions. The system becomes aware of the constructed nature of these precision assignments (e.g., itching directly leading to scratching). Through repeated observation of this process, the system ceases to identify with the precision selections, exactly because it observes them to occur without attributing them to a "self" as an endogenous cause. Recall that the system infers itself as a self—as an endogenous cause of sensations—through agentive actions, where the correspondence between action–outcome contingency feels agentive. Through repeated reorientation of attention back to the meditation target (by reengaging the meditation policy), the automatic precision assignments (e.g., itch→scratch) cease to be identified as essential and so begin to lose the quality of immediateness and irrefutability that come with being transparent. Crucially, as the system observes these precision allocations occurring independently of agentive engagement (due to the non-action policy currently being engaged), the usual means of inferring itself as a self-due to the correspondence of action–outcome contingencies is disrupted, and processes occurring in the system increasingly appear as non-essential to the self. This point about opacification is an important one for our discussion about selfless experiences.

This ties in closely with the "non-self" themes in Buddhism. Common to all Buddhist schools is a critique of our ordinary self-experiences. The Buddhist doctrine of *no-self* (anattā) teaches that our ordinary self-experience is both mistaken and an important source of human suffering. The mistake is the common assumption that behind the various psychophysical processes that make up our conscious experiences, there is a single, essential subject, who is responsible for constructing and owning those processes. However, when we turn our attention inwards in an attempt to catch a glimpse of this assumed subject, all we ever experience is the dynamic and impersonal processes. This is the point—while we commonly assume there to be an essential and *unconstructed* subject, that sense of being a subject is in fact *constructed* from the interaction of our cognitive–behavioral processes. The Buddhist move here is not to deny that the sense of self is real, that quality of experience that delineates my sensations from yours. Rather, the selfless insight Buddhist meditators seek out is the transformative recognition that, while the self unreflectively appears to us as essential, persisting, and unified, it is in fact constructed, impermanent, and dynamic (see Davis and Thompson, 2017).

According to the Buddhist tradition, our mistaken assumptions about the self are generated and maintained by craving (*tanha* in Pali). Craving is a technical term here; it describes the felt urgency or motivational drive to make the world conform to our desires—our anxiety to perpetuate positive feelings and reduce negative ones. This ongoing emotional investment is what unifies the various impersonal psychophysical processes under a single idea: a persistent and essential self. In turn, the more we identify with our specific concerns or roles, the more intense the motivation to bring about those states in the world[5]. In humans, these desires expand beyond basic homeostatic concerns (i.e., being fed and watered) to include the wider constellation of ideas and roles we appropriate into our identity (i.e., being a student). In Buddhism, this craving is thought to be responsible for significant human suffering. Consider the difference in magnitude between the relatively short-lived pain of breaking an arm, and the potentially life long suffering of losing an opportunity to play a beloved sport professionally. At the heart of Buddhism is the teaching that, while pain is unavoidable (i.e., the broken bone), our reaction to it (i.e., the craving to be a professional player) is optional. It is the craving, and not the pain, that is thought to be transformed

[5]For a neuroscientific account of this relationship between identification and motivation, see Damasio (2003).

through meditation and with it the mistaken sense of self that craving engenders.

Meditation is presented as a vehicle for reducing craving and disrupting the mistaken view of self. In the satipatthana sutta, the foremost early Buddhist text on meditation, the student is directed to divide their experience into five categories (or "aggregates") of phenomena: bodily form, valence, perception, volitional activities, and consciousness. These aggregates together are thought to create the experience of being a subject (Hamilton, 2000; Shulman, 2014). The aim of meditation here is to reduce one's identification with these psychophysical processes by closely examining each individually and their interactions and systematically noting their impersonal nature. In Anattalakkhaṇa Suttam, the Buddha suggests that meditators observe each aggregate, saying to themselves "not-I" or "not-my-self." As the various psychophysical processes that make up the self-come to be experienced as "not-I," the constructed nature of the self becomes apparent.

In addition to non-meditation policies (i.e., distractions) being driven by changes in goal states, they are also driven by our affective reactions to those goal states. As we have seen, valence sets precision relative to our ability to reduce error given a certain goal. Ordinarily, once a goal state is selected (predicted), any hesitancy in responding in the ways that the system has learned to expect produces negative valence, which has the effect of increasing the drive on policy selection as an attempt to catch up to the predicted slope. As long as valence is experienced transparently, it has this powerful motivating effect on actions. The meditator then must also be able to reduce the certainty on the valenced reactions that occur from their commitment to non-action. By attending to the valenced sensations themselves (i.e., the discomfort of not reacting), the system learns that these signals too are changing and non-essential. As soon as one has made the observation that non-action makes them uncomfortable, these systems are already being rendered opaque. Focused attention meditation then simultaneously makes opaque precision on goal priors (i.e., the confidence in the degree of wanting or not wanting certain states) and precision on expected free energy reduction that is given affectively.

In Buddhism, reflecting on valence (*vedana* in pali) is considered especially important in the process of relinquishing craving and disrupting the mistaken view of the self. Valence is considered the "weak link" in the process that gives rise to both craving and the mistaken view of self (Anālayo, 2009). In non-meditators, valence conditions craving: pleasant feelings give rise to attachment; painful feelings give rise to resistance. In contrast, long-term meditators are thought to be able to experience valence without further craving-driven responses. The Buddha taught,

> *"Touched by that pleasant feeling he does not lust after pleasure or continue to lust after pleasure. That pleasant feeling of his ceases. With the cessation of the pleasant feeling, painful feeling arises. Touched by that painful feeling, he does not sorrow, grieve, and lament, he does not weep beating his breast and becomes distraught" (Ñānamoli and Bodhi, 1995, p. 334).*

Notice that even for the "well-taught noble disciple" (that is, meditators who no longer operate under craving and illusory notions of the self), valenced states still arise. These states are in and of themselves neutral in terms of well-being (Harris, 2018). It is the strong motivational impulse to act (i.e., fleeing our pain; grasping at joys) in response to those signals that the Buddhist meditative project aims to transform. A close meditative investigation of valence is taught to have the effect of separating valence from craving, the result being that one begins to react with less preference relative to pleasure and pain. As Buddhist scholar Albahari writes, "As mental suffering is finally eliminated through insight [into the non-self-nature of these processes], unpleasant *vedanā* will be confined to only physical (not mental) suffering" (Albahari, 2014, p. 11). As our urgency to react in self-serving ways diminishes, so too does the illusion of being an enduring and essential subject. The culmination of this process of extinguishing craving is *nibbāna* [enlightenment]: "the final flash of insight that burns out *taṇhā* and the sense of self for good" (Albahari, 2014, p. 11).

The gradual opacification of valence results in various positive effects discussed in the Buddhist paradigm. As described above, inaction would allow for the opacification and dereification of the valence system. As valence is increasingly modeled (made opaque), it ceases to invoke that powerful sense of urgency (associated with the Buddhist notion of craving or taṇhaā; Albahari, 2014), which, as we saw above, results from valence being experienced transparently and so presented phenomenologically as both immediately real and essential to the self (Metzinger, 2003). This change is an important one, as it results in the loss of the driving force that perpetuates the sense that there is an essential subject within and behind the various processes (Albahari, 2014). As valence is made opaque, and control over goal prior precisions is achieved, craving ceases due to its disentanglement from the conditioning influence of pleasure or pain. Thus, the illusion of self is disrupted. In gaining endogenous control on the precision of goal priors, meditation therefore enables the system to pull these apart so that (dis)liking does not need to condition (not) wanting, allowing the suffering associated with craving (and aversion) to be avoided. Gaining insight into the process by which valence is driven by our goal priors would have the consequence of allowing one to actively separate the two darts discussed by the Buddha. One comes to understand how pain (deviation from a goal prior) leads to suffering (transparent negative valence signaling the deviation from the goal prior and its expected resolution); that suffering occurs only insofar as we desire to avoid the pain (high precision on the goal prior drives error dynamics); and finally, although we often cannot do anything about the pain itself, we can, by watching the whole process closely, render the valenced reactions opaque and so reduce the degree to which valence drives policy selection. That is, we can reduce the craving and suffering that arises from transparent valenced reactions simply by observing closely the link between pain and discomfort, thereby rendering the valence opaque.

The opacification of this part of the precision machinery opens new opportunities for control. Observing our valenced reactions allows us to develop new higher order policies about

how precision (via valence) is being set on policies[6]. In other words, instead of valenced signals adjusting precision on policies directly, and so automatically conditioning us to behave in certain ways, one can now learn to activate alternative policies depending on the usefulness of the valenced signals. For example, mindfulness has been shown to be highly effective in helping people to quit smoking cigarettes (Bowen and Marlatt, 2009). The practice here is to disrupt the pattern leading from craving to using by selecting a policy to closely attend to the feelings of craving—the negative valence driving processing toward the expected rate of error reduction relative to nicotine levels—every time they arise. The new goal now activated every time there is a craving (instead of smoking a cigarette) is to watch, as closely as possible the arising, the progression and the inevitable depletion of the craving-related feelings. Overtime, these sorts of mindfulness practices have the effect of teaching the system that cravings are in fact just feelings in the body, which can be allowed to direct processing and behavior or not[7]. This discovery can represent a major return of control for people struggling with substance addiction. Notice that valence here does not go missing through this process of opacification (as it does in depersonalization). Rather, valence begins to be interpreted by the system as what it is: information that can be useful, but is not essential, in selecting policies.

CONCLUSION

In this paper, we seek to better understand the nature of the self from the perspective of the increasingly popular active inference framework. Within this framework, we present a novel account of the self, in terms of an allostatic control model (ACM; Deane, 2020). Central to the ACM account is a view of affect as a second-order process that guides the predictive system (via precision weighting) toward opportunities to improve. This is a novel take on affect that we have been developing over a number of recent publications (Kiverstein et al., 2017, 2020; Miller et al., 2020).

We then put this account to work in trying to understand two starkly contrasting forms of self-loss, namely, depersonalization and the selfless experiences attained through meditation (viz., Buddhist no-self insights). Given our proposed framework, these two varieties of selfless experience are characterized by stark differences in the systems degree of control: whereas depersonalization is expressly characterized as resulting from a critical loss of inferred control, selflessness in the context of meditative practices is marked by a significant gain in control.

There is today, however, an increasingly popular idea that depersonalization and the selfless experiences attained through meditation are somehow closely related. Meditation teacher and neuroscience enthusiast Shinzen Young has called depersonalization the "evil twin" of the Buddhist notion of *enlightenment* (Lofthouse, 2014). Part of what motivates this association, so it seems, are similarities in first personal accounts of both kinds of selfless states. Sufferers of depersonalization and long-term meditators make surprisingly similar reports about reductions in their experience of being agents of their actions and as owners of their thoughts and behaviors. While these first personal accounts can sound very similar, given our framework, this is where the family resemblance ends.

As we have shown throughout this paper, there are important computational differences between these two selfless experiences. Of particular importance is the difference in how affective valence contributes to either state. Depersonalization is characterized as a loss of control, leading to the dampening of these affective valence systems. In meditation, there occurs a gradual opacification of the affective valence system. This opacification produces an important change in the meditator's relationship with the positive and negative affect—specifically, by no longer being automatically appropriated into their self-model. This has the result that changes in valence no longer create the existential urgency for change they would otherwise. It is important to note here that valence remains perfectly intact, continuing to tune the agent toward opportunities to improve in their predictive success. The purpose of the practice is not to disrupt valence itself (as happens in depersonalization) but rather to become conscious of the precision estimators in a way that allows them to select more skillful and beneficial policies. What is permanently altered in the meditation process is the system's reaction to those signals. While there is still the experience of frustration and joy, there is no longer the sense of being an essential subject to appropriate these states as "me" and "mine." This insight leads to an eventual dissolving of our misguided idea that we are a single and enduring thing, to be replaced by an acknowledgment that we are a dynamic, self-organizing process. Far from reducing control, and in direct contrast to depersonalization, this development of one's metacognitive abilities here allows one to contextualize and control precision estimations in new and powerful ways.

Notice then that for someone to transition from experiences of depersonalization to the selfless states attained through meditation, they would first need to regain their phenomenal access to those affective responses. Without affect playing its role in tuning the system (relative to its predictive success), the opacification and subsequent insight into the nature of those precision processes could not occur. To be clear, we are not suggesting that the endeavor of bringing affect back online for people suffering depersonalization should be carried out through meditation specifically[8]. Rather, our point is that those affective signals would have to first become available for introspective access in order to be modeled in the way facilitated by focused meditation. This follows from the fact that, on our account, it is not the loss of affectivity that results in the selfless experiences sought after by Buddhist meditators but the process of modeling those affective changes that opens the way for a new perspective of the self and a new layer of control to emerge.

In terms of control, depersonalization and selfless experiences sought after by mediators are, computationally speaking, polar opposites. And yet, there is something important to be said here about the potential focused attention meditation has to

[6]See Smith et al. (2020) for a computational model of this process.

[7]Notice here that the common phenomenology of long-term meditators being able to simply "let go" of mental and emotional distractions is closely related to the *increase* in control that we are proposing this meditative process engenders.

[8]See Lindahl and Britton (2019) for reasons why that might be challenging.

provoke depersonalization experiences. Britton's *Dark Night Project* has documented and investigated a large number of personal reports on various difficulties that can accompany meditative practices[9]. Britton (2019) suggests that adopting a persistent attitude of turning toward difficult stimuli and focusing on negative emotions can lead to negative outcomes. This aspect of mindfulness training has positive effects for some people, primarily by helping them to facilitate a gradual sensitization to negative affect. Such exposure approaches to therapy are thought to work by reducing avoidance, which has been shown to play a leading role in the generation and maintenance of various psychological disorders (Barry et al., 2015). However, exposure therapies are most effective for people who tend toward high levels of avoidance (McNally, 2018). In other, low-avoidance personality types, anxiety and dissociative disorders can be produced and exacerbated by facilitating an attentional bias toward threat (MacLeod et al., 2002; Eldar et al., 2008). The fact is that the most effective treatment for an individual will depend on their baseline attitude toward threat. Given our account above, it makes sense why inappropriately meditating on traumatic events (that is beyond a certain healthy window of tolerance) could create depersonalization effects. If depersonalization is an airbag deployed when fight or flight would not work for getting one out of a traumatic emotional experience, then persistently meditating on an overwhelmingly traumatic experience while practicing inaction produces just those conditions. In effect, it could act like a kind of psychological self-torture. Meditation, in this case, would begin to lead to a perceived loss of allostatic self-efficacy, rather than toward the liberating states of self-understanding it is meant to.

[9] See also Segal (2002) and Lindahl and Britton (2019) for clear accounts of the relationship between meditation and depersonalization.

AUTHOR CONTRIBUTIONS

All authors provided equal contributions to all sections, and so names appear in alphabetical order.

FUNDING

MM and GD carried out this work with the support of Horizon 2020 European Union ERC Advanced Grant XSPECT - DLV-692739.

ACKNOWLEDGMENTS

We would like to thank Lars Sandved-Smith and Maxwell Ramstead for the ongoing discussions about meditation, metacognition, and awareness from an active inference perspective.

REFERENCES

Albahari, M. (2009). Witness-consciousness: its definition, appearance and reality. *J. Conscious. Stud.* 16, 62–84.

Albahari, M. (2014). Insight knowledge of no self in Buddhism: an epistemic analysis. *Philos. Imprint* 14, 1–30.

Allen, M., and Friston, K. J. (2016). From cognitivism to autopoiesis: towards a computational framework for the embodied mind. *Synthese* 195, 2459–2482. doi: 10.1007/s11229-016-1288-5

Allen, M., and Tsakiris, M. (2018). "The body as first prior: interoceptive predictive processing and the primacy," in *The Interoceptive Mind: From Homeostasis to Awareness* (Oxford: Oxford University Press), 2–7.

Anālayo (2009). *From Craving to Liberation-Excursions Into the Thought-word of the Pâli Discourses (1)*. New York, NY: Buddhist Association of the United States.

Apps, M. A., and Tsakiris, M. (2014). The free-energy self: a predictive coding account of self-recognition. *Neurosci. Biobehav. Rev.* 41, 85–97. doi: 10.1016/j.neubiorev.2013.01.029

Ashby, W. (2013). *Design for a Brain: The Origin of Adaptive Behaviour*. Berlin: Springer.

Austin, J. H. (2013). Zen and the brain: mutually illuminating topics. *Front. Psychol.* 4:784. doi: 10.3389/fpsyg.2015.00784

Badcock, P. B., Davey, C. G., Whittle, S., Allen, N. B., and Friston, K. J. (2017). The depressed brain: an evolutionary systems theory. *Trends Cogn. Sci.* 21, 182–194. doi: 10.1016/j.tics.2017.01.005

Baker, D., Earle, M., Medford, N., Sierra, M., Towell, A., and David, A. (2007). Illness perceptions in depersonalization disorder: testing an illness attribution model. *Clin. Psychol. Psychothera. Int. J. Theory Pract.* 14, 105–116. doi: 10.1002/cpp.518

Barrett, L. F. (2017). The theory of constructed emotion: an active inference account of interoception and categorization. *Soc. Cogn. Affect. Neurosci.* 12, 1–23.

Barry, T. J., Vervliet, B., and Hermans, D. (2015). An integrative review of attention biases and their contribution to treatment for anxiety disorders. *Front. Psychol.* 6:968. doi: 10.3389/fpsyg.2015.00968

Barto, A. G. (2013). "Intrinsic motivation and reinforcement learning," in *Intrinsically Motivated Learning in Natural and Artificial Systems*, eds G. Baldassarre and M. Mirolli (Berlin: Springer), 17–47. doi: 10.1007/978-3-642-32375-1_2

Blakemore, S. J., Smith, J., Steel, R., Johnstone, E. C., and Frith, C. D. (2000). The perception of self-produced sensory stimuli in patients with auditory hallucinations and passivity experiences: evidence for a breakdown in self-monitoring. *Psychol. Med.* 30, 1131–1139. doi: 10.1017/s0033291799002676

Bowen, S., and Marlatt, A. (2009). Surfing the urge: brief mindfulness-based intervention for college student smokers. *Psychol. Addict. Behav.* 23:666. doi: 10.1037/a0017127

Britton, W. B. (2019). Can mindfulness be too much of a good thing? The value of a middle way. *Curr. Opin. Psychol.* 28, 159–165. doi: 10.1016/j.copsyc.2018.12.011

Bruineberg, J., and Rietveld, E. (2014). Self-organization, free energy minimization, and optimal grip on a field of affordances. *Front. Hum. Neurosci.* 8:599. doi: 10.3389/fnhum.2014.00599

Carter, O. L., Presti, D. E., Callistemon, C., Ungerer, Y., Liu, G. B., and Pettigrew, J. D. (2005). Meditation alters perceptual rivalry in tibetan buddhist monks. *Curr. Biol.* 15, R412–R413.

Ciaunica, A., Charlton, J., and Farmer, H. (2020). When the window cracks: transparency and the fractured self in depersonalisation. *Phenomenol. Cogn. Sci.* 62, 1–19.

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

Clark, A. (2015). *Surfing Uncertainty: Prediction Action and the Embodied Mind*. Oxford: Oxford University Press.

Colombetti, G., and Ratcliffe, M. (2012). Bodily feeling in depersonalization: a phenomenological account. *Emot. Rev.* 4, 145–150. doi: 10.1177/1754073911430131

Corcoran, A. W., Pezzulo, G., and Hohwy, J. (2019). From allostatic agents to counterfactual cognisers: active inference. *Biol. Philos.* 35:32.

Craig, A. D. (2003). Interoception: the sense of the physiological condition of the body. *Curr. Opin. Neurobiol.* 13, 500–505. doi: 10.1016/s0959-4388(03)00090-4

Damasio, A. (2003). Feelings of emotion and the self. *Ann. N. Y. Acad. Sci.* 1001, 253–261.

Dambrun, M. (2016). When the dissolution of perceived body boundaries elicits happiness: the effect of selflessness induced by a body scan meditation. *Conscious. Cogn.* 46, 89–98. doi: 10.1016/j.concog.2016.09.013

Dambrun, M., and Ricard, M. (2011). Self-centeredness and selflessness: a theory of self-based psychological functioning and its consequences for happiness. *Rev. Gen. Psychol.* 15, 138–157. doi: 10.1037/a0023059

Davis, J., and Thompson, E. (2017). From the five aggregates to phenomenal consciousness: toward a cross-cultural cognitive science. *mindRxiv* [Preprint], Available online at: https://mindrxiv.org/bs8af/

Deane, G. (2020). Dissolving the self: active inference, psychedelics, and ego-dissolution. *Philos. Mind Sci.* 1, 1–21. doi: 10.33735/phimisci.2020.i.39

Eldar, S., Ricon, T., and Bar-Haim, Y. (2008). Plasticity in attention: implications for stress response in children. *Behav. Res. Ther.* 46, 450–461. doi: 10.1016/j.brat.2008.01.012

Fabry, R. E. (2019). Into the dark room: a predictive processing account of major depressive disorder. *Phenomenol. Cogn. Sci.* 19, 685–704. doi: 10.1007/s11097-019-09635-4

Farb, N., Daubenmier, J., Price, C. J., Gard, T., Kerr, C., Dunn, B. D., et al. (2015). Interoception, contemplative practice, and health. *Front. Psychol.* 6:763. doi: 10.3389/fpsyg.2015.00763

Friston, K. (2005). A theory of cortical responses. *Philos. Trans. R. Soc. Lond. B Biol. Sci.* 360, 815–836. doi: 10.1098/rstb.2005.1622

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

Friston, K. (2018). Am I self-conscious?(or does self-organization entail self-consciousness?). *Front. Psychol.* 9:579. doi: 10.3389/fpsyg.2015.00579

Friston, K., FitzGerald, T., Rigoli, F., Schwartenbeck, P., and Pezzulo, G. (2017). Active inference: a process theory. *Neural Comput.* 29, 1–49. doi: 10.1162/neco_a_00912

Friston, K., Rigoli, F., Ognibene, D., Mathys, C., Fitzgerald, T., and Pezzulo, G. (2015). Active inference and epistemic value. *Cogn. Neurosci.* 6, 187–214. doi: 10.1080/17588928.2015.1020053

Friston, K., Schwartenbeck, P., FitzGerald, T., Moutoussis, M., Behrens, T., and Dolan, R. J. (2013). The anatomy of choice: active inference and agency. *Front. Hum. Neurosci.* 7:598. doi: 10.3389/fpsyg.2015.00598

Friston, K. J., and Stephan, K. E. (2007). Free-energy and the brain. *Synthese* 159, 417–458. doi: 10.1007/s11229-007-9237-y

Friston, K. J., Daunizeau, J., Kilner, J., and Kiebel, S. J. (2010). Action and behavior: a free-energy formulation. *Biol. Cybern.* 102, 227–260. doi: 10.1007/s00422-010-0364-z

Friston, K. J., Stephan, K. E., Montague, R., and Dolan, R. J. (2014). Computational psychiatry: the brain as a phantastic organ. *Lancet Psychiatr.* 148–158. doi: 10.1016/s2215-0366(14)70275-5

Frith, C. D. (1987). The positive and negative symptoms of schizophrenia reflect impairments in the perception and initiation of action. *Psychol. Med.* 17, 631–648. doi: 10.1017/s0033291700025873

Frith, C. D. (2014). *The Cognitive Neuropsychology of Schizophrenia*. London: Psychology press.

Garfield, J. L. (2015). "Buddhism and modernity," in *The Buddhist World*, ed. J. Powers (London: Routledge), 94–104.

Fuchs, T. (2005). Corporealized and disembodied minds: a phenomenological view of the body in melancholia and schizophrenia. *Philos. Psychiatr. Psychol.* 12, 95–107.

Gerrans, P. (2019). Depersonalization disorder, affective processing and predictive coding. *Rev. Philos. Psychol.* 10, 401–418. doi: 10.1007/s13164-018-0415-2

Haggard, P. (2017). Sense of agency in the human brain. *Nat. Rev. Neurosci.* 18:196. doi: 10.1038/nrn.2017.14

Hamilton, S. (2000). *Early Buddhism: A New Approach: The I of the Beholder*. Richmond: Curzon Press.

Harris, S. E. (2018). A nirvana that is burning in Hell: pain and flourishing in mahayana buddhist moral thought. *Sophia* 57, 337–347. doi: 10.1007/s11841-017-0619-4

Hesp, C., Smith, R., Allen, M., Friston, K., and Ramstead, M. (2019). Deeply felt affect: the emergence of valence in deep active inference. *PsyArXiv* [Preprint], Available online at: https://psyarxiv.com/62pfd/ (accessed December 3, 2019).

Hohwy, J. (2012). Attention and conscious perception in the hypothesis testing brain. *Front. Psychol.* 3:96. doi: 10.3389/fpsyg.2015.00096

Hohwy, J. (2016). The self-evidencing brain. *Noûs*. 50, 259–285.

Hohwy, J., and Michael, J. (2017). "Why should any body have a self?," in *The Subject's Matter: Self-Consciousness and the Body*, eds A. Amsmith and F. de vignemont (Montréal: MIP Press), 363.

Hosoya, T., Baccus, S. A., and Meister, M. (2005). Dynamic predictive coding by the retina. *Nature* 436, 71–77. doi: 10.1038/nature03689

James, W. (1890). The perception of reality. *Principles Psychol.* 2, 283–324.

Joffily, M., and Coricelli, G. (2013). Emotional valence and the free-energy principle. *PLoS Comput. Biol.* 9:e03094. doi 10.1371/journal.pcbi.1003094

Kaplan, R., and Friston, K. J. (2018). Planning and navigation as active inference. *Biol. Cybernet.* 112, 323–343. doi: 10.1007/s00422-018-0753-2

Kiebel, S. J., Daunizeau, J., and Friston, K. J. (2008). A hierarchy of time-scales and the brain. *PLoS Comput. Biol.* 4:e1000209. doi: 10.1371/journal.pcbi.1000209

Kira, I. A., Ashby, J. S., Odenat, L., and Lewandowsky, L. (2013). The mental health effects of torture trauma and its severity: a replication and extension. *Psychology* 4, 472–482. doi: 10.4236/psych.2013.45067

Kiverstein, J., Miller, M., and Rietveld, E. (2017). The feeling of grip: novelty, error dynamics, and the predictive brain. *Synthese* 196, 2847–2869. doi: 10.1007/s11229-017-1583-9

Kiverstein, J., Miller, M., and Rietveld, E. (2020). How mood tunes prediction: a neurophenomenological account of mood and its disturbance in major depression. *Neurosci. Conscious.* 2020:niaa003. doi: 10.1093/nc/niaa003

Letheby, C., and Gerrans, P. (2017). Self unbound: ego dissolution in psychedelic experience. *Neurosci. Conscious.* 3:nix016. doi: 10.1093/nc/nix016

Limanowski, J., and Blankenburg, F. (2013). Minimal self-models and the free energy principle. *Front. Hum. Neurosci.* 7:547. doi: 10.3389/fpsyg.2015.00547

Lindahl, J. R., and Britton, W. B. (2019). 'I have this feeling of not really being here': buddhist meditation and changes in sense of self. *J. Conscious. Stud.* 26, 157–183.

Linson, A., Parr, T., and Friston, K. J. (2020). Active inference, stressors, and psychological trauma: a neuroethological model of (mal) adaptive explore-exploit dynamics in ecological context. *Behav. Brain Res.* 380:112421. doi: 10.1016/j.bbr.2019.112421

Lofthouse, G. (2014). *Enlightenment's Evil Twin*. Available online at: http://www.theatlantic.com/health/archive/2014/12/enlightenments-eviltwin/383726/ (accessed February, 25 2020).

Lutz, A., Mattout, J., and Pagnoni, G. (2019). The epistemic and pragmatic value of non-action: a predictive coding perspective on meditation. *Curr. Opin. Psychol.* 28, 166–171. doi: 10.1016/j.copsyc.2018.12.019

MacLeod, C., Rutherford, E., Campbell, L., Ebsworthy, G., and Holker, L. (2002). Selective attention and emotional vulnerability: assessing the causal basis of their association through the experimental manipulation of attentional bias. *J. Abnorm. Psychol.* 111:107. doi: 10.1037/0021-843x.111.1.107

Matthews, G. A., and Tye, K. M. (2019). Neural mechanisms of social homeostasis. *Ann. N. Y. Acad. Sci.* 1457, 5–25. doi: 10.1111/nyas.14016

McNally, R. J. (2018). Attentional bias for threat: crisis or opportunity? *Clin. Psychol. Rev.* 69, 4–13. doi: 10.1016/j.cpr.2018.05.005

Medford, N. (2012). Emotion and the unreal self: depersonalization disorder and de-affectualization. *Emot. Rev.* 4, 139–144. doi: 10.1177/1754073911430135

Medford, N., Brierley, B., Brammer, M., Bullmore, E. T., David, A. S., and Phillips, M. L. (2006). Emotional memory in depersonalization disorder: a functional MRI study. *Psychiatry Res. Neuroimag.* 148, 93–102. doi: 10.1016/j.pscychresns.2006.05.007

Medford, N., Sierra, M., Baker, D., and David, A. S. (2005). Understanding and treating depersonalisation disorder. *Adv. Psychiatr. Treat* 11, 92–100. doi: 10.1192/apt.11.2.92

Medford, N., Sierra, M., Stringaris, A., Giampietro, V., Brammer, M. J., and David, A. S. (2016). Emotional experience and awareness of self: functional MRI studies of depersonalization disorder. *Front. Psychol.* 7:432. doi: 10.3389/fpsyg.2015.00432

Metzinger, T. (2003). Phenomenal transparency and cognitive self-reference. *Phenomenol. Cogn. Sci.* 2, 353–393. doi: 10.1023/b:phen.0000007366.42918.eb

Miall, R. C., and Wolpert, D. M. (1996). Forward models for physiological motor control. *Neural Netw.* 9, 1265–1279. doi: 10.1016/s0893-6080(96)00035-4

Miller, M., and Clark, A. (2017). Happily entangled: prediction, emotion, and the embodied mind. *Synthese* 195, 2559–2575. doi: 10.1007/s11229-017-1399-7

Miller, M., Kiverstein, J., and Rietveld, E. K. (2020). Embodying addiction: a predictive processing framework. *Brain Cogn.* 138:105495. doi: 10.1016/j.bandc.2019.105495

Millière, R. (2017). Looking for the self: phenomenology, neurophysiology and philosophical significance of drug-induced ego dissolution. *Front. Hum. Neurosci.* 11:245. doi: 10.3389/fnhum.2016.00245

Millière, R., Carhart-Harris, R. L., Roseman, L., Trautwein, F. M., and Berkovich-Ohana, A. (2018). Psychedelics, meditation, and self-consciousness. *Front. Psychol.* 9:1475. doi: 10.3389/fnhum.2016.001475

Mirza, M. B., Adams, R. A., Mathys, C. D., and Friston, K. J. (2016). Scene construction, visual foraging, and active inference. *Front. Comput. Neurosci.* 10:56. doi: 10.3389/fnhum.2016.00056

Moutoussis, M., Fearon, P., El-Deredy, W., Dolan, R. J., and Friston, K. J. (2014). Bayesian inferences about the self (and others): a review. *Conscious. Cogn.* 25, 67–76. doi: 10.1016/j.concog.2014.01.009

Ñānamoli, B., and Bodhi, B. (1995). *The Middle Length Discourses of the Buddha*. Boston, MA: Wisdom.

Ongaro, G., and Kaptchuk, T. J. (2019). Symptom perception, placebo effects, and the Bayesian brain. *Pain* 160, 1–4. doi: 10.1097/00006396-900000000-98882

Pagnoni, G., and Guareschi, F. T. (2015). Remembrance of things to come: a conversation between Zen and neuroscience on the predictive nature of the mind. *Mindfulness* 8, 27–37. doi: 10.1007/s12671-015-0438-z

Pezzulo, G. (2014). Why do you fear the bogeyman? An embodied predictive coding model of perceptual inference. *Cogn. Affect. Behav. Neurosci.* 14, 902–911. doi: 10.3758/s13415-013-0227-x

Pezzulo, G. (2018). Commentary: the problem of mental action: predictive control without sensory sheets. *Front. Psychol.* 9:1291. doi: 10.3389/fnhum.2016.001291

Pezzulo, G., and Cisek, P. (2016). Navigating the affordance landscape: feedback control as a process model of behavior and cognition. *Trends Cogn. Sci.* 20, 414–424. doi: 10.1016/j.tics.2016.03.013

Pezzulo, G., and Nolfi, S. (2019). Making the environment an informative place: a conceptual analysis of epistemic policies and sensorimotor coordination. *Entropy* 21:350. doi: 10.3390/e21040350

Pezzulo, G., Rigoli, F., and Friston, K. (2015). Active Inference, homeostatic regulation and adaptive behavioural control. *Prog. Neurobiol.* 134, 17–35. doi: 10.1016/j.pneurobio.2015.09.001

Pezzulo, G., Rigoli, F., and Friston, K. J. (2018). Hierarchical active inference: a theory of motivated control. *Trends Cogn. Sci.* 22, 294–306. doi: 10.1016/j.tics.2018.01.009

Radovic, F., and Radovic, S. (2002). Feelings of unreality: a conceptual and phenomenological analysis of the language of depersonalization. *Philos. Psychiatr. Psychol.* 9, 271–279. doi: 10.1353/ppp.2003.0048

Ridderinkhof, K. R., Van Den Wildenberg, W. P., Segalowitz, S. J., and Carter, C. S. (2004). Neurocognitive mechanisms of cognitive control: the role of prefrontal cortex in action selection, response inhibition, performance monitoring, and reward-based learning. *Brain Cogn.* 56, 129–140. doi: 10.1016/j.bandc.2004.09.016

Sass, L. A., and Parnas, J. (2003). Schizophrenia, consciousness, and the self. *Schizophr. Bull.* 29, 427–444. doi: 10.1093/oxfordjournals.schbul.a007017

Schwartenbeck, P., FitzGerald, T., Dolan, R., and Friston, K. (2013). Exploration, novelty, surprise, and free energy minimization. *Front. Psychol.* 4:710. doi: 10.3389/fpsyg.2015.00710

Segal, S. (2002). *Collision with the Infinite: A Life Beyond the Personal Self*. London: New Age Books.

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. K. (2014). *The Cybernetic Bayesian Brain*. Frankfurt: MIND Group.

Seth, A. K., and Friston, K. J. (2016). Active interoceptive inference and the emotional brain. *Philos. Trans. R. Soc. B Biol. Sci.* 371:20160007. doi: 10.1098/rstb.2016.0007

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/fnhum.2016.00395

Seth, A. K., and Tsakiris, M. (2018). Being a beast machine: the somatic basis of selfhood. *Trends Cogn. Sci.* 22, 969–981. doi: 10.1016/j.tics.2018.08.008

Shulman, E. (2014). *Rethinking the Buddha: Early Buddhist Philosophy as Meditative Perception*. Cambridge: Cambridge University Press.

Sierra, M., and Berrios, G. E. (1998). Depersonalization: neurobiological perspectives. *Biol. Psychiatry* 44, 898–908. doi: 10.1016/s0006-3223(98)00015-8

Sierra, M., and David, A. S. (2011). Depersonalization: a selective impairment of self-awareness. *Conscious. Cogn.* 20, 99–108. doi: 10.1016/j.concog.2010.10.018

Sierra, M., Baker, D., Medford, N., and David, A. S. (2005). Unpacking the depersonalization syndrome: an exploratory factor analysis on the Cambridge Depersonalization Scale. *Psychol. Med.* 35, 1523–1532. doi: 10.1017/S0033291705005325

Simeon, D., and Abugel, J. (2006). *Feeling Unreal: Depersonalization Disorder and the Loss of the Self*. Oxford: Oxford University Press.

Simeon, D., Kozin, D. S., Segal, K., Lerch, B., Dujour, R., and Giesbrecht, T. (2008). De-constructing depersonalization: further evidence for symptom clusters. *Psychiatry Res.* 157, 303–306. doi: 10.1016/j.psychres.2007.07.007

Smith, L. S., Hesp, C., Lutz, A., Mattout, J., Friston, K., and Ramstead, M. (2020). Towards a formal neurophenomenology of metacognition: modelling meta-awareness, mental action, and attentional control with deep active inference. *PsyArXiv*. [Preprint]. Available online at: https://psyarxiv.com/5jh3c/

Stephan, K. E., Manjaly, Z. M., Mathys, C. D., Weber, L. A., Paliwal, S., Gard, T., et al. (2016). Allostatic self-efficacy: a metacognitive theory of dyshomeostasis-induced fatigue and depression. *Front. Hum. Neurosci.* 10:550. doi: 10.3389/fnhum.2016.00550

Sterling, P. (2012). Allostasis: a model of predictive regulation. *Physiol. Behav.* 106, 5–15. doi: 10.1016/j.physbeh.2011.06.004

Van de Cruys, S. (2017). *Affective Value in the Predictive Mind*. Frankfurt: MIND Group.

Wilkinson, S., Deane, G., Nave, K., and Clark, A. (2019). "Getting warmer: predictive processing and the nature of emotion," in *The Value of Emotions for Knowledge* (Cham: Palgrave Macmillan), 101–119. doi: 10.1007/978-3-030-15667-1_5

Wolpert, D. M., and Flanagan, J. R. (2001). Motor prediction. *Curr. Biol.* 11, R729–R732.

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: 29 October 2020
doi: 10.3389/fpsyg.2020.523710

Pain Asymbolia as Depersonalization for Pain Experience. An Interoceptive Active Inference Account

*Philip Gerrans**

Department of Philosophy, University of Adelaide, Adelaide, SA, Australia

"Mineness," also called "subjective presence" or "personalization," is the feeling that experiences belong to a continuing self. This article argues that mineness is produced by processes of interoceptive active inference that model the self as the underlying cause of continuity and coherence in affective experience. A key component of this hierarchical processing system and hub of affective self-modeling is activity in the anterior insula cortex. I defend the account by applying it to the phenomenon of pain asymbolia, a condition in which nociceptive signals (of bodily damage) are not attributed to the self. Thus, pain asymbolia is a form of "depersonalization for pain" as Klein puts it. The pain is experienced as happening to *my body* but is not experienced as *mine*. Thus, we can describe it as loss of subjective presence or "mineness" for the experience of pain.

Keywords: interoception, self-awareness, depersonalization, predictive processing, pain asymbolia

OPEN ACCESS

Edited by:
Francesca Ferri,
University of Studies G. d'Annunzio Chieti and Pescara, Italy

Reviewed by:
Anil K. Seth,
University of Sussex, United Kingdom
Vivien Ainley,
University of London, United Kingdom

***Correspondence:**
Philip Gerrans
philip.gerrans@adelaide.edu.au

Specialty section:
This article was submitted to Theoretical and Philosophical Psychology, a section of the journal Frontiers in Psychology

Received: *30 April 2020*
Accepted: *10 September 2020*
Published: *29 October 2020*

Citation:
Gerrans P (2020) Pain Asymbolia as Depersonalization for Pain Experience. An Interoceptive Active Inference Account. Front. Psychol. 11:523710. doi: 10.3389/fpsyg.2020.523710

INTRODUCTION

The topic of this article is a form of experience variously baptized "subjective presence," "mineness," or "personalization" (Seth et al., 2011; Seth, 2013; Billon, 2017a; Guillot, 2017). As the name suggests, it refers to the feeling that experiences belong to a continuing self or comprise autobiographical episodes. The nature and even existence of this elusive phenomenon are contested. However, one important reason for thinking that this form of experience is a genuine phenomenon is a pathological condition in which subjects claim that experiences feel as though they do not "belong to them." In such cases, subjects are not in doubt that they are the subject of experience, sensory, bodily, or cognitive, but they report feeling as though the experience is not "theirs." Such experiences comprise the essence of disorders of depersonalization. These disorders suggest that the feeling of mineness has a distinct phenomenological signature, which can be lost in some conditions, and invites investigation of its causes and typical and atypical manifestation.

As Alexandre Billon puts it:

> *"Every sensation has such a coefficient; we do not notice it, always encountering it; we need to be confronted with exceptional and pathological cases (. . .) to notice it and measure its importance (Billon, 2017a).*

As a way to develop an account of this experience, I focus on a condition called pain asymbolia, aptly described by Klein (2015) as "depersonalization for pain." In pain asymbolia, subjects report feeling detached from painful experience as though it is happening in their body but is not "theirs." It presents as a case of loss of "mineness" for the experience of pain. Pain asymbolia is of particular interest because pain is a bodily state that is normally felt as urgently belonging to the self. For example, Descartes thought it as one of the bodily states that "teach me that. . . I compose a single thing with it [my body]." Wittgenstein's anti-Cartesian meditations on pain were devoted

to explaining the intuition that one could not be in doubt that one was the subject of painful experience. "It is nonsense to say that "I know I am in pain" as it means nothing more than that "I am in pain" (Philosophical Investigations 246).

The account of pain asymbolia I provide situates it in the framework of active inference theories of embodied selfhood, emotion, affect, and self-awareness (Friston et al., 2011; Limanowski and Blankenburg, 2013; Pezzulo et al., 2015; Barrett et al., 2016; Seth and Friston, 2016; Kirchhoff et al., 2018). One version of that framework is proposed by Hohwy and Michael (2017). They argue that experience of embodied selfhood is the product of an inference about the hidden causes of interoceptive (representation of states of the internal *milieu*) experience. On their view, the mind integrates signals from disparate interoceptive channels by inferring that they have a common origin in a unified entity: a bodily self.

This account forms part of an active inference account that treats cognition and action as a hierarchically integrated suite of processes whose goal is reduction of variational free energy. On this account, cognition is the iterative use of generative models (representations) to predict the consequences of actions taken to optimize organismic functioning (Friston, 2010; Hohwy, 2013; Pezzulo et al., 2015). Discrepancies between predicted and actual sensory consequences of action, signaled as prediction errors, entrain the next round of action to reduce error or optimize the model. The active inference theory tells us that prediction error is best minimized over the long term by attributing internally generated interoceptive sensations to a stable, unified entity, a self (Limanowski and Blankenburg, 2013; Hohwy and Michael, 2017; Letheby and Gerrans, 2017). This process of self-modeling creates a basic bodily form of self-awareness.

As Anil Seth puts it:

> *"Mental representations of selfhood are ultimately grounded in representations of the body, with the internal physiological milieu providing a primary reference—a '**material me**'" (Seth, 2014).*

As suggested by Seth, *material me* provides an anchor for other forms of self-representation. Sensorimotor control, agency, perceptual perspective, and explicit narrative self-representation are cognitive processes that require a form of implicit or explicit self-representation: a model of the entity that sustains the relevant process. The basic sense of being the continuing subject of experience on whose behalf all these activities are performed is the awareness of material me and underpins other forms of self-representation and awareness.

The explanation of mineness and its absence in depersonalization I propose focuses attention on a particular form of interoceptive self-representation, namely, *affective* self-representation. In affective self-representation, the mind models the bodily self as the source of affective experience and target of affective regulation. This level of self-modeling produces the experience of being the person/entity whose affective states modulate as her goals are realized or frustrated in action. We can call this the experience of being ***affective me***.

At still higher levels of self-modeling, we deploy explicit conceptual or imagistic representations of the self. This is the level at which we self-attribute character and personality traits using a self-concept. We can call this self-model **narrative me** because it models the self as the protagonist of a recountable autobiography (Schechtman, 2011; Goldie, 2011).

Affective me and interoceptive me are very closely related, because affective experience is a form of interoceptive experience. However, they are distinguishable, particularly in conditions such as depersonalization in which they dissociate. These conditions do not seem to be the result of loss of interoceptive or other basic capacities for body sensation and regulation. Rather, I shall argue that depersonalization is the result of a failure of affective self-modeling resulting from impairment in the neural substrates of affective me. Someone with depersonalization experience has the intractable experience of an intact material and narrative me, combined with hypoactivity in the circuitry that sustains affective me.

The concept of an affective me produced by interoceptive active inference integrates and synthesizes ideas advanced in different forms by a variety of theorists across disciplines of psychiatry neuroscience cognitive science and philosophy. I first explain the concept of interoceptive active inference and show how it explains (i) how interoception and affective experience are related via the process of hierarchical self-modeling, (ii) how the framework explains the role of the anterior insula cortex (AIC) in producing the experience of "mineness," "subjective presence," or "personalization" of experience. I then apply the framework to the explanation of pain aysmbolia. Pain asymbolia is a case in which nociceptive signals (of bodily damage) are not integrated with affective signals because of hypoactivity in the anterior insula. The mind, however, *predicts* that affective me will feel distress as a consequence of pain/nociception. The result is a prediction error that cannot be resolved because relevant affective and self-modeling mechanisms are deactivated. As a result, when an experience of pain, predicted to have a strong affective signature, does not produce affect, the subject feels as though it is not happening to her.

MATERIAL ME: INTEROCEPTION AS ALLOSTATIC ACTIVE INFERENCE

Interoceptive and affective states have a common basis in allostatic (action to optimize internal body states in context) regulation. Allostasis is a refinement of the concept of homeostasis, which implies a reflexive return to an optimal "set point" for levels of basic bodily function such as blood oxygenation. Allostasis extends that concept, recognizing that optimality for some variables requires variation according to context rather than maintenance of a single optimal set point. In fact, it can be helpful to think of homeostasis and allostasis as on a continuum of flexibility. Some functions (like blood oxygenation) have very tight parameters and are context insensitive. Others (like blood pressure) need to fluctuate more widely to sustain viability of the organism. Thus, some variables representing body state are monitored not only in relation to a homeostatic set point, but also in relation to their departure from a level predicted as optimal in context (Barrett and Simmons, 2015;

Barrett et al., 2016; Corcoran and Hohwy, 2017; Kleckner et al., 2017). Allostasis thus introduces an element of forecasting to homeostatic regulation.

> *"Interoception and homeostatic regulation are inevitably linked and form a closed loop: tuning the set points of homeostatic reflex arcs depends on accurate allostatic predictions about future bodily states; these predictions, in turn, depend on accurate inference about current bodily states" (Stephan et al., 2016).*

Interoception is the integrated representation of information about states of very basic, dynamically controlled, bodily processes such as blood oxygenation and endocrine and electrolyte balance for the purpose of allostatic regulation. In order to regulate the body, interoception models the hidden causes of allostatic fluctuations by attributing them to a unified entity (Limanowski and Blankenburg, 2013; Moutoussis et al., 2014; Sel, 2014; Barrett, 2017; Seth and Tsakiris, 2018; Wiese, 2018). The unified hidden cause of allostatic variation along multiple dimensions tracked and integrated in interoception is material me.

Interoception effectively integrates disparate streams of information about basic bodily regulation to inform us of *global organismic state* relative to predicted state. As Seth and Tsakiris (2018) point out, we feel the results of dehydration, poisoning or deoxygenation, but the effects are felt globally at the level of conscious awareness in sensations such as fatigue. When we attend to states such as thirst or fatigue, we do not succeed in more precisely representing the causal structure of the entities responsible for the experience (for example, the molecular mechanisms of dehydration or shifts in the production of metabolites and effects on neurotransmission). Rather, the goal is to establish the degree of departure from optimality of a global feeling state so that we can manage it at the systemic level (e.g., by drinking or resting). Interoceptive experience thus provides a personal-level proxy for the regulation of low-level homeostatic/allostatic variables whose mechanisms are opaque to introspection Joffily and Coricelli, 2013). The predictive or forecasting aspect is introduced by the need to regulate, by anticipating interoceptive fluctuations, and to evaluate actual state against those predictions. In other words, when we feel fatigue, we feel overall energy depletion relative to a prediction of optimal energy levels for that context. Interoceptive regulation uses experience to predict

> *"how an action would affect physiological homeostasis, given a model" (Seth and Tsakiris, 2018).*

The idea that interoception is experienced as a systemic non-localized phenomenon connects with an interesting metaphysical point made by Wiese (2018) in his discussion of self-representation and predictive coding. Like Letheby and Gerrans (2017) and Hohwy and Michael, he argues that, phenomenologically, the self *seems* to be a substance: an enduring entity that underlies changing sensations and perceptions. Wiese points out, however, that interoceptive experience does not specify a particular localizable entity in a way that allows further discovery by the attentive deployment of perception or theoretical inference, because there is no concrete (by which he means spatiotemporally located) object of interoception to focus on. Wiese has an intricate and sophisticated predictive processing account of the sense of being the subject of experience as an "abstract enduring object" to which experience is *salient*. *Abstract* because it is a higher-order amodal integrator of lower-order information streams, and *enduring* because it represents the continuing entity in which those streams cohere.

> *"the apparent substantiality of the phenomenal self is explained by a structural feature of this salience model: it binds different dimensions of salience by representations of higher-order dimensions of salience (just as more abstract object representations bind representations of perceptual features in predictive processing accounts of feature binding)" (Wiese, 2018).*

The similarity between material me, *qua* object of interoception, and objects of perception is their explanatory role as underlying hidden cause of coherence in experience. The difference between material me and objects of visual perception is the non-concreteness of material me. Seth and Tsakiris make a similar point to Weise.

> *"instrumental (control-oriented) interoceptive inference plausibly underlies a phenomenology related to the evaluation of the allostatic consequences of regulatory actions. A non-localized, non–object-based phenomenology associated with both mood and emotion, and with the pre-reflective (i.e., non-reflexive) self-related experience of being an embodied organism (Seth and Tsakiris, 2018).*

AFFECTIVE ME: EMOTION AS INTEROCEPTIVE ACTIVE INFERENCE

So far, we have only explained why interoception creates "the pre-reflective self-related experience of being an embodied organism" as Seth and Tsakiris put it. The short answer is that allostatic regulation requires us to experience ourselves in interoception as an integrated entity to serve as the target of regulation. However, we have not begun to explain (i) how it is that interoception is associated with mood and emotion and (ii) how it is that this account can be mobilized to explain how people can feel detached from their bodily experiences in cases of depersonalization and pain asymbolia.

To do so, we need to explain the higher levels and dimensions of interoceptive self-modeling. The starting point is to note that some signals of body state are "vital signs." Fatigue, sustained high temperature, or intractable nociception threatens the organism and requires urgent action. Consequently, we have evolved the capacity to feel such states, not simply as perturbations of body state, but as urgently motivating. Affective processes provide this "feeling of what matters" to slightly modify Antonio Damasio's phrase.

Damasio's account is a neo (William) Jamesian account of emotion, affect, and self-awareness that grounds all these experiences in bodily processing. The history of this idea and the way it is expressed are not uniform across the disciplines. Not everyone is using terms the same way. So, somewhat stipulatively, let me say I am using the term *self-awareness* to refer to a *feeling* of being the entity whose continued life underpins other

forms of experience. I use the term *representation* to refer to information-bearing structures. There is no implication that the content of such structures is always consciously experienced. Thus, self-representation can be conscious or unconscious, but self-awareness is experienced. And I use emotion in a standard way derived from analysis of prototypical episodes of mental life (for example sadness, nostalgia, anger) that are evaluative and motivational, have characteristic bodily and behavioral indices, and are, typically, felt (Deonna and Teroni, 2012). I use the term *affect* to refer to a form of experience that carries emotional content. Thus, affect is common to emotion, which usually has an identifiable eliciting object, and mood, which does not. The feeling of anxiety can be part of an episode of emotion (anxiety about a specific forthcoming event) or, in the case of anxious mood, a feeling of hypervigilance and uncertainty without a particular object.

The idea that emotional episodes have evaluative, behavioral, cognitive, and affective components is part of the appraisal theory of emotion (Scherer, 2004; Grandjean et al., 2008). On this theory, emotional processes evaluate (appraise) the relevance of events (including internal events such as allostatic prediction error) for the organism. We might say that cognition and perception represent aspects of the world, and emotional processes represent the significance of that information for the well-being of the organism. And to do so, they need to model the organism as an entity with goals realized or frustrated in action (including internal regulatory action) (Scherer, 2004; Kalisch et al., 2006; Grandjean et al., 2008; Kalisch, 2009; Brosch and Sander, 2013).

This explains the subtle relationship between affective processing and interoception. Interoception integrates and aggregates allostatic variables to inform us of global organismic state. Emotional processes evaluate interoceptive signals against expectations about goal satisfaction in context. The result is experienced as an affective state. And affective states inherit from the interoceptive processes they metarepresent two interesting properties. They are intimately felt as states of a self, and at same time, they are global and non-localized. We do not experience sadness as a change in the state of a perceived object but as a global, overall feeling state of affective me. And in the same way as interoceptive experience is a proxy for allostatic regulation affective experience is a higher-level proxy for lower-level regulation. Experience or anticipation of danger, for example, prompts a suite of cognitive and behavioral responses that entrain a set of lower-level activities designed to optimize organismic function.

In order to provide the affective interpretation of changes in body state, affective processes exploit extra layers of emotional processing that metarepresent and interpret interoceptive signals (Stephan et al., 2016). The emotional interpretation of interoceptive signals requires integrating the interoceptive signal with information about the emotional salience of the situation and the subject's affective history. In other words, "should material me expect to feel like this given the (emotional) context?" For predictive processing theories of emotion, predictive models of the emotional context set parameters that determine how physiological changes are regulated and experienced. Thus, interdependent models representing the emotional world (hostile, favorable, tractable *for me*) and the capacities of the organism to deal with that world interpret and predict interoceptive changes in a continuous cycle. It is in that sense that emotional processes are forms of interoceptive active inference (Barrett and Simmons, 2015; Barrett et al., 2016; Barrett, 2017; Kleckner et al., 2017). They provide higher-level interpretive and regulatory models for the reduction of interoceptive prediction error signals. Thus, emotional processes are part of a hierarchy of active inference. Interoception is allostatic active inference, and emotion is interoceptive active inference.

Still higher levels of self-modeling, narrative or conceptual, interpret and predict states of affective me, and one can see that other forms of self-representation and related experience, agential, and sensorimotor are guided and reinforced by their effects on affective me. In the end, without affective me, we are an organism to whom the world and its own states, as represented by our battery of cognitive faculties, no longer matter.

NEURAL CORRELATES OF MATERIAL AND AFFECTIVE ME: EMOTIONAL TRANSCRIPTION

Affective experience is produced by emotional processes that integrate interoceptive information with perception and cognition to produce the "feeling of what happens," in Antonio Damasio's phrase. To create this feeling, emotional processes effectively transcribe bodily feelings into affective experiences. This is why an interoceptive state such as fatigue can be experienced as disconsolate apathy when transcribed emotionally as part of an episode of sadness. As a bodily state, it has a particular experiential signature. Transcribed by emotional processes as the state of a self, rendered hopeless by an irretrievable loss, it has an affective signature. Given the way emotional and interoceptive processes are woven together, almost every interoceptive state is transcribed like this. It takes disorders and dissociations, such as pain asymbolia or rare states of emotional neutrality, to decompose their interactions. There is an analogy with delusions of misidentification (DMS) based on loss of predicted affective response to familiar faces. Normally, the feeling of familiarity evoked by recognizing a face is not salient amid the flux of experience, but when it is absent, the experience is of seeing a familiar but feeling "as if" one sees a stranger. Similarly, when one's own body is damaged, the mind predicts an affective response. When that feeling is absent, it feels "as if" the experience is not happening to the subject. It is actually a striking parallel between disorders of depersonalization and the experiential (or first stage as it is sometimes called) component of DMS that they are reported in "as if" vocabulary (Breen et al., 2001; Brighetti et al., 2007; Coltheart et al., 2010). In both cases, the loss of predicted affective response combined with preserved cognition in other relevant domains (face recognition or nociception/interoception) creates the experience of estrangement.

The fatigue of depression and its morphing into disconsolate apathy provides a nice case study of the role of hierarchical

interoceptive inference in producing affective me. Fatigue is an adaptive state designed to restore depleted subsystemic function. If, however, fatigue is intractably sustained, the result will be a persistent homeostatic/allostatic error signal experienced as a characteristic interoceptive state: weariness and exhaustion. The active inference hierarchy will exploit higher-level models to interpret and contextualize these interoceptive signals (Barrett et al., 2016; Friston et al., 2018; Velasco and Loev, 2020). Those higher-level models include models of affective me and narrative me that predict the effects on feeling state of activity across system, given the subject's life history. Initially, such models predict restoration consequent on rest. If, however, the homeostatic error signals "from below" cannot be canceled, higher-level models can be revised to reflect that signal. This reflects the general principle that the hierarchical processing system settles into a state that minimizes error across the system. The self will be modeled as unable to control basic states and to act efficaciously in the world. At the level of affective me, the state is now felt as apathy and anhedonia, possibly anxiety at the prospect of exertion. At the level of narrative self-representation, thoughts of hopelessness and inadequacy can come to dominate. As Stephan et al., put it in their predictive processing account of fatigue and depression:

> *"belief of failure at one's most fundamental task—homeostatic/allostatic regulation—. . . arises from experiencing enhanced interoceptive surprise. We suggest that fatigue is a (possibly adaptive) initial allostatic response to a state of interoceptive surprise; if dyshomeostasis continues, the belief of low allostatic self- efficacy and lack of control may pervade all domains of cognition and manifests as a generalized sense of helplessness, with depression as a consequence" (Stephan et al., 2016).*

This quotation suggests a bottom-up etiology ("belief arises"), but active inference accounts such as this also give a constitutive role to top-down models in reconfiguring low-level allostatic processing. As Seth puts it:

> *"On this theory of interoceptive inference [. . .], emotional states (i.e., subjective feeling states) arise from top-down predictive inference of the causes of interoceptive sensory signals [. . .]" (Seth and Friston, 2016, p. 9).*

The hierarchical nature of interoceptive processing and emotional processing is reflected in cytoarchitecture (Barrett and Simmons, 2015). Sections of the posterior insula cortex (PIC) take primary interoceptive afferents and integrate those representations to coordinate basic regulatory functions. The PIC, for example, integrates values of allostatic variables such as blood pressure and hydration as well as nociception (bodily damage). These signals are progressively remapped and integrated with other information at higher levels of cognition. Although it is not the only channel for bodily signaling, its role as a primary integrative hub of interoceptive afferents makes PIC a crucial substrate of the experience of material me (Singer et al., 2009; Medford and Critchley, 2010; Gasquoine, 2014; Moayedi, 2014).

The AIC is specialized to re-represent and integrate information about body state to allow us to feel the significance of interoceptive states as affects. AIC sits at the apex of the so-called "salience system," the neural hierarchy that signals whether and how information matters to the organism. In order to perform its role, it must communicate with emotional processing hubs that coordinate appraisal of that information at all levels (Craig, 2009a; Craig, 2009b; Garfinkel and Critchley, 2013). The AIC is an enigmatic and functionally ubiquitous system widely connected to both lower- and higher-level processing circuitry. One telling feature is its consistent involvement in the processing of self-relevant information and the switching/anticorrelation between executive and default processing (a key substrate of narrative me) (Starr et al., 2009). Another is its connectivity to hubs of lower- and higher-level emotional appraisal (amygdala and ventromedial prefrontal cortex). These hubs coordinate the appraisal of and response to information presented in sensorimotor and higher-level cognition, respectively (Bechara et al., 1999; Scherer, 2004; Koenigs and Grafman, 2009; Adolphs, 2010; Pessoa and Adolphs, 2010; Gerrans and Scherer, 2013; LeDoux and Pine, 2016). Thus, the AIC is activated by perception of emotionally salient/self-relevant scenarios. Its activity allows us to feel the emotional significance of events. It is also active in scenarios in which we reflect on past events or anticipate the future, allowing so-called mental time travel to be imbued with affective significance. Thus, AIC activity can also provide the affective texture for anticipation and recollection as well as sensory processing. When imagining going on holiday or getting married, the affective texture of the episode predicts how the action will make us feel. This is a crucial adaptation for learning and planning. We plan our actions on behalf of future states of affective me. To do so, we remember how previous actions made us feel. This way of conceptualizing mental time travel makes it a form of extended active inference in which affective regulation is a proxy for overall organismic regulation (Suddendorf and Corballis, 2007; Boyer, 2008; Buckner et al., 2008; Broyd et al., 2009; Spreng et al., 2009; Carhart-Harris and Friston, 2010).

Thus, it is not surprising to see that contemporary affective neuroscience treats experience produced by AIC activation as a form of higher-order bodily representation that represents the integrated functioning of the organism *evaluated against emotionally salient goals creating a sense of self in the process* As Bud Craig (2009a) puts it:

> *"The integration successively includes homeostatic, environmental, hedonic, motivational, social, and cognitive activity to produce a 'global emotional moment,' which represents the sentient self at one moment of time" (Craig, 2009a).*

In other words, affective me is the body under an emotional mode of presentation (to import some philosophical jargon). I am endorsing Craig's idea that the AIC plays a key role in transforming interoceptive signals integrated by the PIC into representations of states of an affective self. This process can be described as one of higher-order metarepresentation and interpretation but to do so underplays constitutive influence of higher-order affective processing on lower-order processing. The role of emotional processing is not just to determine the relevance of interoceptive predictive error *after the fact* but also to set the parameters that determine which allostatic variations become

prediction errors. How the bodily signal is processed depends on how it is emotionally contextualized.

Thus, I prefer to describe the relationship between AIC and PIC as emotional transcription. The AIC is an integrative hub of processes that transform a bodily into an affective signal, in the process transforming what would otherwise be a pure bodily feeling into an affective/emotional one. The AIC communicates with hubs of emotional processes to convert the neural signal representing interoceptive information into a neural signal representing the emotional significance of that information for the organism.

This transcription creates the crucial affective dimension of self-modeling. Without it, we could navigate the world using other dimensions, agential, narrative, sensorimotor, and bodily in order to optimize organismic functioning, however, we would lack a way to experience the *significance* of our interactions with the world.

Furthermore, this affective dimension of self-modeling provides a simple and effective proxy for the adaptive integration of other dimensions and the systems they coordinate. When our narrative, agential sensorimotor, agential, and bodily self-models are simultaneously optimized, we feel good: our organism is prospering in the world. When it is not, we feel a form of negative affect appropriate to the context. Within the predictive coding framework, negative affect signals failure to reduce prediction error across the system. More accurately, it reflects failure to reduce prediction at a rate predicted for that action (Joffily and Coricelli, 2013).

This is not a claim that the AIC is a discrete or modularized substrate of self-representation. Rather, its integrative role, connecting low- and high-level emotional processing and interoception and in the generation of affective experience, makes it an important hub of processing that enables us to feel the significance of events as affective states. Predictive coding suggests that the mind will model fluctuations in affective states by attributing them to a continuing entity: the thing that experiences the emotional ups and downs. Affective me is that entity. And a good candidate for its neural substrate, given its role as a hub of self-referential processing, is the AIC.

PAIN ASYMBOLIA AS A FORM OF DEPERSONALIZATION EXPERIENCE

The idea that the AIC is the substrate of affective self-modeling fits with studies of (relatively) selective damage or hypoactivity of AIC in disorders of depersonalization. In depersonalization disorder, patients report phenomenology such as the following:

> *"I feel some degree of 'out of it' all the time (. . .) I can sit looking at my foot or my hand and not feel like they are mine. This can happen when I am writing, my hand is just writing, but I'm not telling it to. It almost feels like I have died, but no one has thought to tell me. So, I'm left living in a shell that I don't recognize any more" (Sierra and David, 2011).*

Within the multidimensional multilevel framework, this could be explained as a result of preserved narrative/conceptual levels of self-modeling and basic bodily self-modeling in the absence of the basic experience of *being* the self who experiences autobiographical episodes. Another classic description of generalized depersonalization is Dugas' patient who said:

> *"I only feel anger from the outside, by its physiological reactions" (Dugas and Moutier, 1911) my italics," quoted in Billon (2017b).*

This is a particularly telling example because it suggests that bodily processes and bodily awareness are intact, but the patient feels detached from them, despite awareness that they occur in his own body. On the account developed above, that is the result of failure to transcribe bodily interoceptive signals into affective.

The mechanism is sometimes hypothesized to be spontaneous inhibition of AIC by the ventrolateral prefrontal cortex (Medford et al., 2006; Medford, 2012; Medford et al., 2016). On most accounts, this is an involuntary defensive/dissociative response to unmanageable adversity. Of course, given the integrative role of the AIC and its dense multidirectional coupling and functional connectivity, there may be no unique cause of hypoactivity. What matters to the account here is that in the experience of depersonalization AIC hypoactivity is unpredicted by a self-model that anticipates AIC activity in context. It could be the case that the AIC is not receiving interoceptive afferents from lower levels or systems that appraise those signals for emotional relevance. If, however, the AIC is not responding in a predictable way to those afferents, the result is an error signal experienced as loss of predicted affect. The result is that most dimensions of self-representation, bodily, agential sensorimotor, and narrative are intact, but the agent does not feel as if any of the resultant experiences belong to her. The reason is that the AIC is no longer functioning to allow her to feel the significance of bodily changes evoked by her passage through the world (Gerrans, 2015, 2019). It is worth mentioning here that this interpretation of the role of the AIC has been disputed on the bases of cases of AIC lesion with "preserved emotional and affective responses" (Philippi et al., 2012; Damasio et al., 2013; Feinstein et al., 2016). My reading of these cases, however, is that the patients have intact *behavioral* aspects of emotion (such as aversive response) and primary interoception, which accounts for intact bodily feeling. However, the same patients do not seem to exhibit affective aspects of emotion such as feelings of sadness or remorse. Similarly, their empathic responses are cognitive rather than affective. In fact, the profile of Roger, the subject of discussion in two key articles, somewhat resembles that of Dugas' patient (Gerrans, 2019). Very interestingly, Roger cannot be aversively conditioned to painful stimuli, although he responds aversively on each separate presentation is intact. This suggests that he does not anticipate negative affective experience when re-presented with the aversive stimulus.

Rather than discuss full-blown global depersonalization, this section concentrates on a fascinating subtype of depersonalization experience in which only one channel of processing is disconnected from affective me. Pain asymbolia is in which the subject feels pain or its nociceptive aspects, but says that the pain feels as if it does not matter or does not belong to her.

Colin Klein has argued persuasively that pain asymbolia is a form of depersonalization for pain. As he puts it:

> *"the phenomenology of asymbolia might resemble a kind of depersonalization syndrome. . . . The asymbolic, and the depersonalized more generally, feel sensations that they are estranged from—that they do not take to be theirs in the sense that we normally do. . . . [This]does show that there is another sense in which our sensations may be unified: as sensations over which we have a feeling of ownership. Asymbolia, and depersonalization more generally, shows that this sort of unity may fail. Its failure comes not from a change in the sensations we feel, but in the sort of agents we are (Klein, 2015) [my italics].*

Klein suggests that the pain sensation is unchanged, but what has changed in the experience is "the sort of agents we are," i.e., the type of self we are. This can be finessed still further when we add that this latter change must itself be experienced. Otherwise, the patient would not report the feeling that the pain sensation does not matter to *her*. Given the previous discussion, it is not quite right to say that the sensations have not changed, but the experience of the self has. Given that the self-model sets the parameters for bodily representation and consequent experience, a change in self-model affects the quality of experience.

Pain asymbolia is a nice example of the connection between basic bodily processing, emotional processing, and affect. Pain itself is a representation of damaged body state (nociception), but given its significance for organisms, the nociceptive signal is almost automatically appraised at primary level as distressing. Thus, pain, aversive response, and negative affect are very tightly linked (Krahé et al., 2013; Klein, 2015; Gogolla, 2017; Von Mohr and Fotopoulou, 2018; Gehrlach et al., 2019). Another way to put this is to say that bodily damage is represented at multiple levels in terms of its effects not just on the body but on the self and its prospects.

> *"Pain can therefore constitute a process of perceptual inference about nociceptive signals on the basis of predictive, top-down signals about the homeostatic significance of such signals in the context of other synchronous biological, cognitive, and social conditions. Furthermore, such re-mappings of interoceptive signals across the neurocognitive hierarchy suggest possible neurobiological mechanisms by which not only cognitive, but also social contextual factors can influence the awareness of interoceptive and other multimodal information about one's own body" (Krahé et al., 2013).*

Given these facts, the substrate of pain experience is a complex network of nociceptive, interoceptive, and social/cognitive/emotional circuitry. This "pain matrix" incorporates the insula as well as somatosensory and limbic regions (Starr et al., 2009). Characterizing the role of the insula in the matrix is complex due to its extensive connectivity, but one study reports a consensus view:

> *"insula may be well positioned to utilize cognitive information to modulate connected brain areas involved in processing of sensory-discriminative, affective, and cognitive-evaluative components of pain" (Wiech and Tracey, 2013).*

or, as I might put it, to help coordinate higher-level active inference in response to lower-level nociceptive prediction error. And one feature of this coordinating role is the production of affective states that inform the organism of the significance of bodily damage. When, however (due to hypoactivity in the AIC), the predicted negative affect does not occur, the subject has to explain away the resultant prediction error. The bodily self-model is already functioning optimally: telling the organism that she is damaged. The narrative model is also intact: it says explicitly that pain should produce negative affect. Thus, the patient is in the situation of sensing bodily damage and knowing, intellectually, that she has bodily damage but feeling no distress in a situation in which she normally feels it automatically. She reports the result as the feeling that the pain is not "hers." Pain asymbolics no longer assign or feel emotional significance in response to bodily damage in virtue of hypoactivity in their AIC. Effectively, they are in the situation of losing a crucial dimension of affective self-modeling for nociception. Consequently, they report that the experience is painful but that it does not matter and feels as if it is not happening to *them*. What this shows is that "mineness" can be lost locally, for aspects of bodily functioning, such as pain (Phillips et al., 2001; Phillips and Sierra, 2003; Medford et al., 2006; Simeon and Abugel, 2006; Simeon et al., 2008; Stein and Simeon, 2009; Sierra et al., 2012; Michal et al., 2013; Sedeño et al., 2014; Medford et al., 2016; Gogolla, 2017; Gerrans, 2019).

The idea that the anterior insula is a substrate of the feeling of mineness for pain via its role in affective processing is consistent with the similarity between depersonalization experience for pain and mild opioid analgesia. In opioid analgesia, patients report that the pain is not extinguished but *no longer matters*. A key finding here is that that opioids target not only the PIC, as one might expect, but also the AIC and related limbic structures involved in emotional processing.

The AIC in fact is even more responsive than PIC to low doses of opioids. This is an adaptation. It is easier for an organism to regulate emotional/affective response to bodily damage than to repair bodily damage. Thus, in contexts where the organism cannot devote resources to repair, it inhibits the system that produces negative affect and thereby stops pain from drawing attention away from other relevant activities. Opioids exploit this adaptation, down-regulating the AIC, reducing, not pain itself, but the felt significance of pain.

> *"the FMRI data suggest that opioid analgesics can directly influence emotional responses at low doses that do not alter sensory aspects of pain" (Lee et al., 2014).*

Another way to put this is to say that mild opioid analgesia produces a mild form of pain asymbolia.

This suggestion about the role played by modulation of AIC activity independent of nociception is supported by an interesting finding about *voluntary imagination of sensory states*.

> *"For the visualization of internal state sensations, this meant increased activity in areas of interoceptive sensory processing, including the mid and anterior insula in the right hemisphere. This is a critical finding, as it suggests that primary interoceptive cortex, located in the posterior insula, was not significantly involved in the imagery of internal state sensations" (Bennett and Baird, 2009).*

Cases like this suggest that when we imagine or reflect on an experience, we (re)construct the affective component of experience. In other words, we represent not what happens to us but how it matters to us. We empathize with our past or future self by activating circuitry, which represents not body state *per se*, but the significance of body state. This allows us to enrich the narrative self-model with episodic and affective imagery, transforming it from a linguistic autobiographical model to one which we feel, as well as know, is ours. In other words, activity in the AIC links the affective dimension of self-modeling to our bodily and narrative models. In this respect, the affective self-model has a crucial integrative role. It sits at the border between sensorimotor bodily modeling, which controls organismic interaction with the environment, and explicit top-down cognition, which exploits narrative and conceptual models. It allows us to feel not just like a cognitive system manipulating a body through the world (which we are) but a self, with an autobiographical trajectory that matters.

CONCLUSION

The mind models and predicts fluctuations of affect by attributing them to a continuing self. That "self-model" allows us to experience not just the way things are, but the way they matter to us given our history, goals, and concerns.

Perhaps the most crucial dimension of self-modeling is affective. The ability to feel the significance of our engagements with the world allows us to regulate our organism moment to moment and offline over long time scales to "feel the future" and rehearse the past. In this sense, affective processes knit systemic functioning together allowing us to pursue organismic well-being by regulating our affective states.

When this integrative process fails (due to hypoactivity in the AIC or systems that link affective processing to cognition), but the world and body are being otherwise accurately represented, the subject feels that something is wrong. Furthermore, her autobiographical knowledge is undisturbed. The result is a massive prediction error in the hierarchical multidimensional self-model. The narrative and bodily dimensions are intact, but the predicted affective dimension is absent. She reports the result naturally enough by saying that the experience feels as if it is not happening to her.

DATA AVAILABILITY STATEMENT

All datasets generated for this study are included in the article/supplementary material, further inquiries can be directed to the corresponding author/s.

AUTHOR CONTRIBUTIONS

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

REFERENCES

Adolphs, R. (2010). What does the amygdala contribute to social cognition? *Ann. N. Y. Acad. Sci.* 1191, 42–61. doi: 10.1111/j.1749-6632.2010.05445.x

Barrett, L. F. (2017). The theory of constructed emotion: an active inference account of interoception and categorization. *Soc. Cogn. Affect. Neurosci.* 12, 1–23.

Barrett, L. F., Quigley, K. S., and Hamilton, P. (2016). An active inference theory of allostasis and interoception in depression. *Phil. Trans. R. Soc. B* 371:20160011. doi: 10.1098/rstb.2016.0011

Barrett, L. F., and Simmons, W. K. (2015). Interoceptive predictions in the brain. *Nat. Rev. Neurosci.* 16, 419–429. doi: 10.1038/nrn3950

Bechara, A., Damasio, H., Damasio, A. R., and Lee, G. P. (1999). Different contributions of the human amygdala and ventromedial prefrontal cortex to decision-making. *J. Neurosci.* 19, 5473–5481. doi: 10.1523/jneurosci.19-13-05473.1999

Bennett, C. M., and Baird, A. A. (2009). The processing of internally-generated interoceptive sensation. *Neuroimage* 47, S84.

Billon, A. (2017a). *Mineness first, The Subjects Matter: Self-Consciousness and the Body*. Cambridge, MA: MIT Press.

Billon, A. (2017b). "What Is it like to lack mineness? Depersonalization as a probe for the scope, nature and role of mineness," in *The Sense of Mineness* (Cambridge MA: OUP).

Boyer, P. (2008). Evolutionary economics of mental time travel? *Trends Cogn. Sci.* 12, 219–224. doi: 10.1016/j.tics.2008.03.003

Breen, N., Coltheart, M., Caine, D., Lewis, M. B., and Ellis, H. D. (2001). A two-way window on face recognition. *Trends Cogn. Sci.* 5, 234–235. doi: 10.1016/s1364-6613(00)01659-4

Brighetti, G., Bonifacci, P., Borlimi, R., and Ottaviani, C. (2007). "Far from the heart far from the eye": evidence from the Capgras delusion. *Cogn. Neuropsychiatry* 12, 189–197. doi: 10.1080/13546800600892183

Brosch, T., and Sander, D. (2013). Comment: the appraising brain: towards a neuro-cognitive model of appraisal processes in emotion. *Emot. Rev.* 5, 163–168. doi: 10.1177/1754073912468298

Broyd, S. J., Demanuele, C., Debener, S., Helps, S. K., James, C. J., and Sonuga-Barke, E. J. (2009). Default-mode brain dysfunction in mental disorders: a systematic review. *Neurosci. Biobehav. Rev.* 33, 279–296. doi: 10.1016/j.neubiorev.2008.09.002

Buckner, R. L., Andrews-Hanna, J. R., and Schacter, D. L. (2008). The brain's default network: anatomy, function, and relevance to disease. *Ann. N. Y. Acad. Sci.* 1124, 1–38. doi: 10.1196/annals.1440.011

Carhart-Harris, R. L., and Friston, K. J. (2010). The default-mode, ego-functions and free-energy: a neurobiological account of freudian ideas. *Brain* 133, 1265–1283. doi: 10.1093/brain/awq010

Coltheart, M., Menzies, P., and Sutton, J. (2010). Abductive inference and delusional belief. *Cogn. Neuropsychiatry* 15, 261–287. doi: 10.1080/13546800903439120

Corcoran, A. W., and Hohwy, J. (2017). Allostasis, interoception, and the free energy principle: feeling our way forward. *PsyArXiv [Preprint]* doi: 10.31234/osf.io/zbqnx.

Craig, A. (2009a). Emotional moments across time: a possible neural basis for time perception in the anterior insula. *Phil Trans. Biol. Sci.* 364, 1933–1942. doi: 10.1098/rstb.2009.0008

Craig, A. D. (2009b). How do you feel–now? the anterior insula and human awareness. *Nat. Rev. Neurosci.* 10, 59–70. doi: 10.1038/nrn2555

Damasio, A., Damasio, H., and Tranel, D. (2013). Persistence of feelings and sentience after bilateral damage of the insula. *Cereb. Cortex* 23, 833–846. doi: 10.1093/cercor/bhs077

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

Feinstein, J. S., Khalsa, S. S., Salomons, T. V., Prkachin, K. M., Frey-Law, L. A., Lee, J. E., et al. (2016). Preserved emotional awareness of pain in a patient with extensive bilateral damage to the insula, anterior cingulate, and

amygdala. *Brain Struct. Funct.* 221, 1499–1511. doi: 10.1007/s00429-014-0986-3

Dugas, L., and Moutier, F. (1911). *La dépersonnalisation. F. Alcan.* Available online at: http://www.biusante.parisdescartes.fr/histmed/medica/cote?79749

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

Friston, K., Mattout, J., and Kilner, J. (2011). Action understanding and active inference. *Biol. Cybern.* 104, 137–160. doi: 10.1007/s00422-011-0424-z

Friston, K. J., Joffily, M., and Barrett, L. F. (2018). "Active inference and emotion," in *The Nature of Emotion: Fundamental Questions*, 2nd Edn, eds A. S. Fox, R. C. Lapate, A. J. Shackman, and R. J. Davidson (New York, NY: Oxford University Press), 28–33.

Garfinkel, S. N., and Critchley, H. D. (2013). Interoception, emotion and brain: new insights link internal physiology to social behaviour. Commentary on:"Anterior insular cortex mediates bodily sensibility and social anxiety" by Terasawa et al.(2012). *Soc. Cogn. Affect. Neurosci.* 8, 231–234. doi: 10.1093/scan/nss140

Gasquoine, P. G. (2014). Contributions of the insula to cognition and emotion. *Neuropsychol. Rev.* 24, 77–87. doi: 10.1007/s11065-014-9246-9

Gehrlach, D. A., Dolensek, N., Klein, A. S., Chowdhury, R. R., Matthys, A., Junghänel, M., et al. (2019). Aversive state processing in the posterior insular cortex. *Nat. Neurosci.* 22, 1424–1437. doi: 10.1038/s41593-019-0469-1

Gerrans, P. (2015). "Metamisery and bodily inexistence," in *Open MIND* T. Metzinger and J. M. Windt eds (Frankfurt am Main: MIND Group)

Gerrans, P. (2019). Depersonalization disorder, affective processing and predictive coding. *Rev. Phil. Psychol.* 10, 401–418. doi: 10.1007/s13164-018-0415-2

Gerrans, P., and Scherer, K. (2013). Wired for despair the neurochemistry of emotion and the phenomenology of depression. *J. Conscious. Stud.* 20, 254–268.

Gogolla, N. (2017). The insular cortex. *Curr. Biol.* 27, R580–R586.

Goldie, P. (2011). "Life, fiction, and narrative," in *Narrative, Emotion, and Insight*, eds N. Carroll and J. Gibson (Pennsylvania: Pennsylvania State University Press), 8–22.

Grandjean, D., Sander, D., and Scherer, K. R. (2008). Conscious emotional experience emerges as a function of multilevel, appraisal-driven response synchronization. *Conscious. Cogn.* 17, 484–495. doi: 10.1016/j.concog.2008.03.019

Guillot, M. (2017). I me mine: on a confusion concerning the subjective character of experience. *Rev. Phil. Psychol.* 8, 23–53. doi: 10.1007/s13164-016-0313-4

Hohwy, J. (2013). *The Predictive Mind*. Oxford: Oxford University Press.

Hohwy, J., and Michael, J. (2017). "16 Why should any body have a self?," in *The Subject's Matter: Self-Consciousness and the Body*, eds F. de Vignemont, A. Alsmith (Cambridge, MA: The MIT Press) 363.

Joffily, M., and Coricelli, G. (2013). Emotional valence and the free-energy principle. *PLoS Comput. Biol.* 9:e1003094. doi: 10.1371/journal.pcbi.1003094

Kalisch, R. (2009). The functional neuroanatomy of reappraisal: time matters. *Neurosci. Biobehav. Rev* 33, 1215–1226. doi: 10.1016/j.neubiorev.2009.06.003

Kalisch, R., Wiech, K., Critchley, H. D., and Dolan, R. J. (2006). Levels of appraisal: a medial prefrontal role in high-level appraisal of emotional material. *Neuroimage* 30, 1458–1466. doi: 10.1016/j.neuroimage.2005.11.011

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

Kleckner, I. R., Zhang, J., Touroutoglou, A., Chanes, L., Xia, C., Simmons, W. K., et al. (2017). Evidence for a large-scale brain system supporting allostasis and interoception in humans. *Nat. Hum. Behav.* 1:0069.

Klein, C. (2015). What pain asymbolia really shows. *Mind* 124, 493–516. doi: 10.1093/mind/fzu185

Koenigs, M., and Grafman, J. (2009). The functional neuroanatomy of depression: distinct roles for ventromedial and dorsolateral prefrontal cortex. *Behav. Brain Res.* 201, 239–243. doi: 10.1016/j.bbr.2009.03.004

Krahé, C., Springer, A., Weinman, J. A., and Fotopoulou, A. K. (2013). The social modulation of pain: others as predictive signals of salience–a systematic review. *Front. Hum. Neurosci.* 7:386. doi: 10.3389/fnhum.2013.00386

LeDoux, J. E., and Pine, D. S. (2016). Using neuroscience to help understand fear and anxiety: a two-system framework. *Am. J. Psychiatry* 173, 1083–1093. doi: 10.1176/appi.ajp.2016.16030353

Lee, M. C., Wanigasekera, V., and Tracey, I. (2014). Imaging opioid analgesia in the human brain and its potential relevance for understanding opioid use in chronic pain. *Neuropharmacology* 84, 123–130. doi: 10.1016/j.neuropharm.2013.06.035

Letheby, C., and Gerrans, P. (2017). Self unbound: ego dissolution in psychedelic experience. *Neurosci. Conscious.* 2017, nix016.

Limanowski, J., and Blankenburg, F. (2013). Minimal self-models and the free energy principle. *Front. Hum. Neurosci.* 7:547. doi: 10.3389/fnhum.2013.00547

Medford, N. (2012). Emotion and the unreal self: depersonalization disorder and de-affectualization. *Emot. Rev.* 4, 139–144. doi: 10.1177/1754073911430135

Medford, N., Brierley, B., Brammer, M., Bullmore, E. T., David, A. S., and Phillips, M. L. (2006). Emotional memory in depersonalization disorder: a functional MRI study. *Psychiatry Res. Neuroimaging* 148, 93–102. doi: 10.1016/j.pscychresns.2006.05.007

Medford, N., and Critchley, H. D. (2010). Conjoint activity of anterior insular and anterior cingulate cortex: awareness and response. *Brain Struct. Funct.* 214, 535–549. doi: 10.1007/s00429-010-0265-x

Medford, N., Sierra, M., Stringaris, A., Giampietro, V., Brammer, M. J., and David, A. S. (2016). Emotional experience and awareness of self: functional MRI studies of depersonalization disorder. *Front. Psychol.* 7:432. doi: 10.3389/fpsyg.2016.00432

Michal, M., Koechel, A., Canterino, M., Adler, J., Reiner, I., Vossel, G., et al. (2013). Depersonalization disorder: disconnection of cognitive evaluation from autonomic responses to emotional stimuli. *PLoS One* 8:e74331. doi: 10.1371/journal.pone.0074331

Moayedi, M. (2014). All roads lead to the insula. *Pain* 155, 1920–1921. doi: 10.1016/j.pain.2014.07.023

Moutoussis, M., Fearon, P., El-Deredy, W., Dolan, R. J., and Friston, K. J. (2014). Bayesian inferences about the self (and others): a review. *Conscious. Cogn.* 25, 67–76. doi: 10.1016/j.concog.2014.01.009

Pessoa, L., and Adolphs, R. (2010). Emotion processing and the amygdala: from a'low road'to'many roads' of evaluating biological significance. *Nat. Rev. Neurosci.* 11, 773–783. doi: 10.1038/nrn2920

Pezzulo, G., Rigoli, F., and Friston, K. (2015). Active inference, homeostatic regulation and adaptive behavioural control. *Prog. Neurobiol.* 134, 17–35. doi: 10.1016/j.pneurobio.2015.09.001

Philippi, C. L., Feinstein, J. S., Khalsa, S. S., Damasio, A., Tranel, D., Landini, G., et al. (2012). Preserved self-awareness following extensive bilateral brain damage to the insula, anterior cingulate, and medial prefrontal cortices. *PLoS One* 7:e38413. doi: 10.1371/journal.pone.0038413

Phillips, M. L., Medford, N., Senior, C., Bullmore, E. T., Suckling, J., Brammer, M. J., et al. (2001). Depersonalization disorder: thinking without feeling. *Psychiatry Res. Neuroimaging* 108, 145–160. doi: 10.1016/s0925-4927(01)00119-6

Phillips, M. L., and Sierra, M. (2003). Depersonalization disorder: a functional neuroanatomical perspective. *Stress* 6, 157–165. doi: 10.1080/1025389031000138538

Schechtman, M. (2011). "The narrative self," in *The Oxford Handbook of the Self*, ed. Gallagher S. (Oxford: Oxford University Press), 394–416.

Scherer, K. R. (2004). "Feelings integrate the central representation of appraisal-driven response organization in emotion," in *Feelings and Emotions: the Amsterdam Symposium*, eds A. S. R. Manstead, N. Frijda, and A. Fischer (Cambridge: Cambridge University Press), 136–157. doi: 10.1017/cbo9780511806582.009

Sedeño, L., Couto, B., Melloni, M., Canales-Johnson, A., Yoris, A., Baez, S., et al. (2014). How do you feel when you can't feel your body? Interoception, functional connectivity and emotional processing in depersonalization-derealization disorder. *PLoS One* 9:e98769. doi: 10.1371/journal.pone.0098769

Sel, A. (2014). Predictive codes of interoception, emotion, and the self. *Front. Psychol.* 5:189. doi: 10.3389/fpsyg.2014.00189

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. K. (2014). "The cybernetic bayesian brain," in *Open MIND*, eds T. K. Metzinger and J. M. Windt (Frankfurt am Main: MIND Group).

Seth, A. K., and Friston, K. J. (2016). Active interoceptive inference and the emotional brain. *Phil. Trans. R. Soc. B Biol. Sci.* 371:20160007. doi: 10.1098/rstb.2016.0007

Seth, A. K., Suzuki, K., and Critchley, H. D. (2011). An interoceptive predictive coding model of conscious presence. *Front. Psychol.* 2:395. doi: 10.3389/fpsyg.2011.00395

Seth, A. K., and Tsakiris, M. (2018). Being a beast machine: the somatic basis of selfhood. *Trends Cogn. Sci.* 22, 969–981. doi: 10.1016/j.tics.2018.08.008

Sierra, M., and David, A. S. (2011). Depersonalization: a selective impairment of self-awareness. *Conscious. Cogn.* 20, 99–108. doi: 10.1016/j.concog.2010.10.018

Sierra, M., Medford, N., Wyatt, G., and David, A. S. (2012). Depersonalization disorder and anxiety: a special relationship? *Psychiatry Res.* 197, 123–127. doi: 10.1016/j.psychres.2011.12.017

Simeon, D., and Abugel, J. (2006). *Feeling Unreal: Depersonalization Disorder and the Loss of the Self*. Oxford: Oxford University Press.

Simeon, D., Kozin, D. S., Segal, K., Lerch, B., Dujour, R., and Giesbrecht, T. (2008). De-constructing depersonalization: further evidence for symptom clusters. *Psychiatry Res.* 157, 303–306. doi: 10.1016/j.psychres.2007.07.007

Singer, T., Critchley, H. D., and Preuschoff, K. (2009). A common role of insula in feelings, empathy and uncertainty. *Trends Cogn. Sci.* 13, 334–340. doi: 10.1016/j.tics.2009.05.001

Spreng, R. N., Mar, R. A., and Kim, A. S. (2009). The common neural basis of autobiographical memory, prospection, navigation, theory of mind, and the default mode: a quantitative meta-analysis. *J. Cogn. Neurosci.* 21, 489–510. doi: 10.1162/jocn.2008.21029

Starr, C. J., Sawaki, L., Wittenberg, G. F., Burdette, J. H., Oshiro, Y., Quevedo, A. S., et al. (2009). Roles of the insular cortex in the modulation of pain: insights from brain lesions. *J. Neurosci.* 29, 2684–2694. doi: 10.1523/jneurosci.5173-08.2009

Stein, D. J., and Simeon, D. (2009). Cognitive-affective neuroscience of depersonalization. *CNS Spectr.* 14, 467–471. doi: 10.1017/s109285290002352x

Stephan, K. E., Manjaly, Z. M., Mathys, C. D., Weber, L. A., Paliwal, S., Gard, T., et al. (2016). Allostatic self-efficacy: a metacognitive theory of dyshomeostasis-induced fatigue and depression. *Front. Hum. Neurosci.* 10:550.

Suddendorf, T., and Corballis, M. C. (2007). The evolution of foresight: What is mental time travel, and is it unique to humans? *Behav. Brain Sci.* 30, 299–313. doi: 10.1017/s0140525x07001975

Velasco, P. F., and Loev, S. (2020). Affective experience in the predictive mind: a review and new integrative account. *Synthese* 1–36.

Von Mohr, M., and Fotopoulou, A. (2018). "The cutaneous borders of interoception: active and social inference of pain and pleasure on pain and pleasure on the skin," in *The Interoceptive Mind: From Homeostasis to Awareness* eds M. Tsakiris, and H. De Preester (Oxford: Oxford University Press), 102.

Wiech, K., and Tracey, I. (2013). Pain, decisions, and actions: a motivational perspective. *Front. Neurosci.* 7:46. doi: 10.3389/fnins.2013.00046

Wiese, W. (2018). *Experienced Wholeness. Integrating Insights from Gestalt Theory, Cognitive Neuroscience, and Predictive Processing*. Cambridge, MA: MIT Press.

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 in Psychology

BRIEF RESEARCH REPORT
published: 05 November 2020
doi: 10.3389/fpsyg.2020.531567

"Dance Like Nobody's Watching": Exploring the Role of Dance-Based Interventions in Perceived Well-Being and Bodily Awareness in People With Parkinson's

Rebecca Hadley[1], Olivia Eastwood-Gray[1], Meryl Kiddier[2], Dawn Rose[3] and Sonia Ponzo[1]*

[1] Department of Psychology and Sports Sciences, University of Hertfordshire, Hatfield, United Kingdom, [2] Ready Steady Move, Chelmsford, United Kingdom, [3] School of Music, Lucerne University of Applied Sciences and Arts, Lucerne, Switzerland

Evidence indicates that bodily perception is negatively related to Parkinson's disease (PD); in particular, people with Parkinson's (PwP) feel dissatisfaction in their physical abilities and appearance. While established treatments exist to ameliorate motor symptoms in PD, research has yet to explore the effects of well-being-focused interventions in relation to the subjective experience of bodily concerns of PwP. This mixed methods exploratory study investigated the constructs of body appreciation in relation to well-being in PwP and the impact of participation in a dance class on body appreciation and well-being, comparing PwP with age-matched controls. Participants (n = 27 PwP, n = 14 controls) completed the Warwick Edinburgh Mental Well-Being Scale and the Body Appreciation Scale before and after taking part in a dance class. Well-being was positively associated with body appreciation in PwP (r_S = 0.64, $p < 0.001$) but not in controls. Following participation in a dance class, all participants' well-being scores increased; a greater increase in well-being scores was observed for controls. A pilot qualitative study explored bodily awareness with PwP who attended dance classes (n = 4) and other movement-based activities (n = 4). Analysis of the interview data indicated that PwP who danced showed heightened bodily awareness, including bodily limitations, in comparison with PwP who did not dance. These preliminary findings provide initial insight explaining the lack of improvements in body appreciation in PwP following a dance class. The current study highlights the need for dance interventions for PwP to consider incorporating elements that encourage a body positive attitude alongside fostering perceived well-being.

Keywords: Parkinson's disease, bodily awareness, dance, body appreciation, well-being, exercise

OPEN ACCESS

Edited by:
Anna Ciaunica,
University of Porto, Portugal

Reviewed by:
Sabine Glock,
University of Wuppertal, Germany
Mateusz Wozniak,
Central European University, Hungary

****Correspondence:***
Rebecca Hadley
r.hadley3@herts.ac.uk;
rebeccahadley3@gmail.com

Specialty section:
This article was submitted to Theoretical and Philosophical Psychology, a section of the journal Frontiers in Psychology

Received: *31 January 2020*
Accepted: *22 September 2020*
Published: *05 November 2020*

Citation:
Hadley R, Eastwood-Gray O, Kiddier M, Rose D and Ponzo S (2020) "Dance Like Nobody's Watching": Exploring the Role of Dance-Based Interventions in Perceived Well-Being and Bodily Awareness in People With Parkinson's. Front. Psychol. 11:531567. doi: 10.3389/fpsyg.2020.531567

INTRODUCTION

Parkinson's disease (PD) is a chronic neurodegenerative condition with an approximated worldwide prevalence rate of 425 per 100,000 persons aged 65–74 years (Pringsheim et al., 2014). The number of individuals living with the condition globally is estimated to have doubled between 1990 and 2016 to over 6 million (Dorsey et al., 2018). Clinical presentation of motor symptoms

typically includes slowness of movement, muscular rigidity, and/or tremor (Kalia and Lang, 2015). Non-motor difficulties also present in PD include neuropsychiatric conditions (e.g., anxiety and depression), sleep disorders, fatigue, and pain (Chaudhuri et al., 2011). Even in the early stages, PwP report that non-motor symptoms have a greater impact on health-related quality of life than motor features (Pfeiffer, 2016). As the severity of the condition progresses, worsening of both motor and non-motor symptoms may affect an individual's ability to undertake essential daily tasks, such as washing and dressing (Duncan et al., 2014). The complex interplay between motor and non-motor symptoms leads to reduced physical activity (Nimwegen et al., 2011) and multifaceted stressors on psychological well-being (Nicoletti et al., 2017) for PwP.

The majority of studies that have focused on non-motor symptoms (e.g., depression; Okun and Watts, 2002, general well-being; Jenkinson et al., 1995) has not considered the relationship between PD and bodily perception as a mediating factor (Koller, 1984). Existing evidence indicates a relationship between PD and bodily perception (Koller, 1984) as well as dissatisfaction toward one's own body's physical abilities and appearance (Koller, 1984; **Gamarra et al., 2009**). Accordingly, neuroimaging findings (see Christopher et al., 2014 for a meta-analysis) suggest that bodily concerns in PD may be linked to damage to the insular cortex, a key region involved in bodily awareness (namely, awareness of one's own body; Craig, 2002). Despite growing evidence, little is known about the impact of bodily related concerns on PwP quality of life. Only a few studies (e.g., Ghielen et al., 2017) propose integrated approaches to increase PwP well-being while taking into account bodily awareness.

In other conditions, including fibromyalgia, chronic fatigue syndrome (Courtois et al., 2015), post-traumatic stress disorder (Mehling et al., 2018), and anorexia nervosa (Kolnes, 2017), exercise has been used as a therapeutic strategy that can enhance bodily awareness and improve well-being. Dance, in particular, is associated with a heightened connectedness to the body as dancers communicate through performing physical movement (Walter and Yanko, 2018). For example, contemporary dance has been associated with higher body appreciation (i.e., satisfaction with and acceptance of one's own body appearances and abilities; Avalos et al., 2005; Tylka and Wood-Barcalow, 2015) in dancers compared to non-dancers (Swami and Tovée, 2009). Furthermore, taking part in dance has been found to improve self-reported positive well-being in young adults (Kim and Kim, 2007) and reduce depressive symptoms in older adults living in a nursing home (Vankova et al., 2014). For PwP, dance classes have been recognized as being useful in the management of both the physical and psychological components of well-being (Reynolds et al., 2016). Specifically, dance has been shown to increase the well-being of attendees due to its high level of engagement and unique social components (Shanahan et al., 2015). Quantitative studies have demonstrated significant meaningful improvement in quality of life for PwP after regular participation in tango dance classes (Hackney and Earhart, 2009). An increase in psychological well-being has also been reported for PwP after attending a single dance session (Koch et al., 2016). Qualitative research has indicated that the social value of dancing with a partner, or as a group, fosters relationships and reduces feelings of loneliness (Houston and McGill, 2013; Holmes and Hackney, 2017). However, few studies have considered how dancing affects body appreciation and whether this construct may be related to perceived well-being in PwP. A more comprehensive assessment of such mediating factors, i.e., bodily awareness and appreciation, may help health professionals understand more about potential barriers to participation and inform their practice.

Previous studies have suggested that dance, as a body-focused activity, may benefit perceived well-being in PwP, although the role that bodily awareness and appreciation might play has not been studied in relation to PwP. Therefore, the present study explored the relationship between perceived well-being and body appreciation, as well as the effect of dance classes on both constructs in PwP and controls. Specifically, the aims of the current study were to investigate quantitatively (a) whether the constructs of well-being and body appreciation were correlated with one another in PwP compared to controls and (b) whether taking part in a dance class shapes well-being and body appreciation and specifically so in PwP compared with age-matched controls. We predicted (a) that self-reported well-being would be positively associated with body appreciation for PwP and for controls and (b) higher body appreciation and well-being for both PwP and age-matched controls post-dance class compared to pre-scores. Semi-structured interviews were undertaken as part of a pilot qualitative study to gain further insight into how bodily awareness is experienced by PwP who attend dance classes and PwP who attend other types of movement-based activities, excluding dance.

QUANTITATIVE STUDY

Methods

Participants

A total of 41 people, 27 with Parkinson's (16 female, age range = 47–86 years, $M = 69.04$, $SD = 8.56$ months, average age at diagnosis = 63 years) and 14 controls (10 female, age range = 52–81 years, $M = 68.57$, $SD = 8.72$), were recruited via dance classes specifically devised for PwP throughout Hertfordshire and the surrounding counties. Controls ($n = 14$) were the partners or caregivers of the PwP and attended the same classes. Exclusion criteria was a diagnosis of a neurological disorder other than Parkinson's but did not include psychiatric disorders due to the comorbidity with PD (e.g., depression; Schapira et al., 2017). The majority of participants (PwP and controls) were physically active each week (**Supplementary Table 1**) and reported dance as one of the activities they participated in (PwP; $n = 23$ and controls; $n = 12$). Most participants with PD ($n = 26$) were taking dopaminergic medication. This study was approved by the Health, Sciences, Engineering and Technology Ethics Committee with Delegated Authority at the University of Hertfordshire [Protocol reference: aLMS/UG/UH/03538(3) and aLMS UG UH 03539(1)]. All participants gave informed, written consent.

Materials

Two dependent variables were collected; well-being was assessed using the Warwick–Edinburgh Mental Well-Being Scale (WEMWBS; Tennant et al., 2007) and positive body image assessed using the Body Appreciation Scale-2 (BAS-2; Tylka and Wood-Barcalow, 2015).

The WEMWBS is a self-report measure capturing the hedonistic and eudemonic elements of positive mental well-being (Tennant et al., 2007). It includes 14 items such as, "I've been feeling useful" and "I've been feeling good about myself." A 5-point Likert scale (1 = none of the time and 5 = all of the time) is provided for each item. The minimum score is 14, and the maximum is 70. The overall score is calculated by summing the scores across all items. A higher score indicates higher mental well-being. The WEMWBS has been found to have a high internal reliability for a general population sample (Cronbach's $\alpha = 0.91$; Tennant et al., 2007).

The BAS-2 is a self-report measure of an individual's acceptance and positive perceptions of their body. It includes 10 items such as, "I respect my body." and "I am attentive to my body's needs." Each item is rated on a 5-point scale (1 = never and 5 = always). The minimum score is 10, and the maximum score is 50. The total score is calculated by summing the scores across all items. A higher score indicates greater body appreciation. The BAS-2 has been found to have a high internal reliability for a general population sample (Cronbach's $\alpha = 0.97$; Tylka and Wood-Barcalow, 2015).

Experimental Procedure

Demographic information collected included symptoms of Parkinson's, current medications, and types of physical activity regularly undertaken.

Participants (control and PwP) completed the WEMWBS and BAS-2 (both in written form) immediately before (pre) and after (post) taking part in a singular "Dance for Parkinson's" class. For the few PwP who had difficulty completing the handwritten forms, either their partner or a volunteer helped record their responses.

Each dance class was an hour long and began with a 10–15 min seated warm-up. To accommodate a range of Parkinson's severities, the sessions (30–40 min) included dance steps that could be performed to music while seated or standing. Unlike other studies that have used a specific dance style (i.e., Tango Argentino; Hackney and Earhart, 2009; Koch et al., 2016; Beerenbrock et al., 2020), the current study used a genre classified as social dance (Lewis et al., 2016). This required participants to perform solo and partnered routines of a variety of styles (i.e., Bollywood, Tango, Irish dance). The last 10 min of the class involved a seated cool-down.

Data Analysis

Non-parametric statistical tests were employed for data analysis due to the small sample size of the control group and the non-normal distribution of WEMWBS and BAS-2 scales (see **Figures 3**, **4** below). To test whether wellbeing and body appreciation improved for PwP following participation in a dance class (compared with controls), main effects of (i) group (PwP vs Controls) and (ii) time (pre vs post) were calculated. The main effect of group was calculated by averaging the pre and post WEMWBS and BAS-2 scores for the PD and control group separately. Values were then compared between groups using a Mann-Whitney test. The main effect of time was calculated by averaging the values of pre and post WEMWBS and BAS-2 scores, regardless of group classification. These values were compared via a Wilcoxon signed-rank test. To assess the presence of two-way interactions between group and time, post-dance-values of WEMWBS and BAS-2 were subtracted from pre-dance values and differences between the two groups analyzed using Mann–Whitney *U*-tests. *Post hoc* tests (Bonferroni corrected $\alpha = 0.025$) were employed to investigate significant two-way interactions (tested via Wilcoxon signed-rank tests). Effect sizes were computed using Cohen's $r(r = Z/\sqrt{N})$. Spearman correlations between WEMWBS and BAS-2 baseline (pre) scores in (i) the Parkinson's sample and (ii) the controls were conducted to assess the relationship between the two constructs irrespective of time.

One control participant (female, aged 75 years) did not complete the WEMWBS, and one control (female, aged 69 years) and one PwP (male, aged 72 years) did not complete the BAS-2. These data were not included in the analyses.

Data were analyzed and plotted using R (Allen et al., 2019).

Results

Correlations

To investigate whether the constructs of well-being and body appreciation are related, only the pre-intervention scores were used, and their relationship was explored via Spearman correlations (two-tailed) in the Parkinson's sample as well as in controls. Pre-intervention data, unaffected by the experimental manipulation, were used, as they reflect the stable state of these traits.

A significant positive correlation was found between well-being scores (as measured via the WEMWBS) and body appreciation scores (BAS-2) in the Parkinson's sample [$r_s(24) = 0.64$, $p < 0.001$; **Figure 1**]. However, no significant correlation was found between well-being and body appreciation scores in the control sample [$r_s(10) = 0.07$, $p = 0.42$; **Figure 2**]. This finding suggests that, for PwP specifically, increases in body appreciation are associated with increases in perceived well-being, but this is not the case for controls. However, the lack of effect in the control group may be due to the insufficient experimental power, given the small sample size.

Well-Being

To investigate whether participation in a dance class shapes perceived well-being (measured using the WEMWBS) in PwP compared with age-matched controls, Mann-Whitney U-tests were used to analyze between-subjects factor of a 2 (groups: PwP vs. control) × 2 (time: pre vs. post) mixed design. Within-subjects factors were analyzed via Wilcoxon signed-rank tests. A significant main effect of time overall ($Z = -4.02$, $p < 0.001$, $r = 0.64$) revealed that post-dance scores were higher than pre-dance ones (pre: *median* = 47.50, *interquartile range* = 12.50, $M = 46.83$, $SD = 11.78$; post: *median* = 51.50, *interquartile range* = 18.00, $M = 51.05$, $SD = 12.45$).

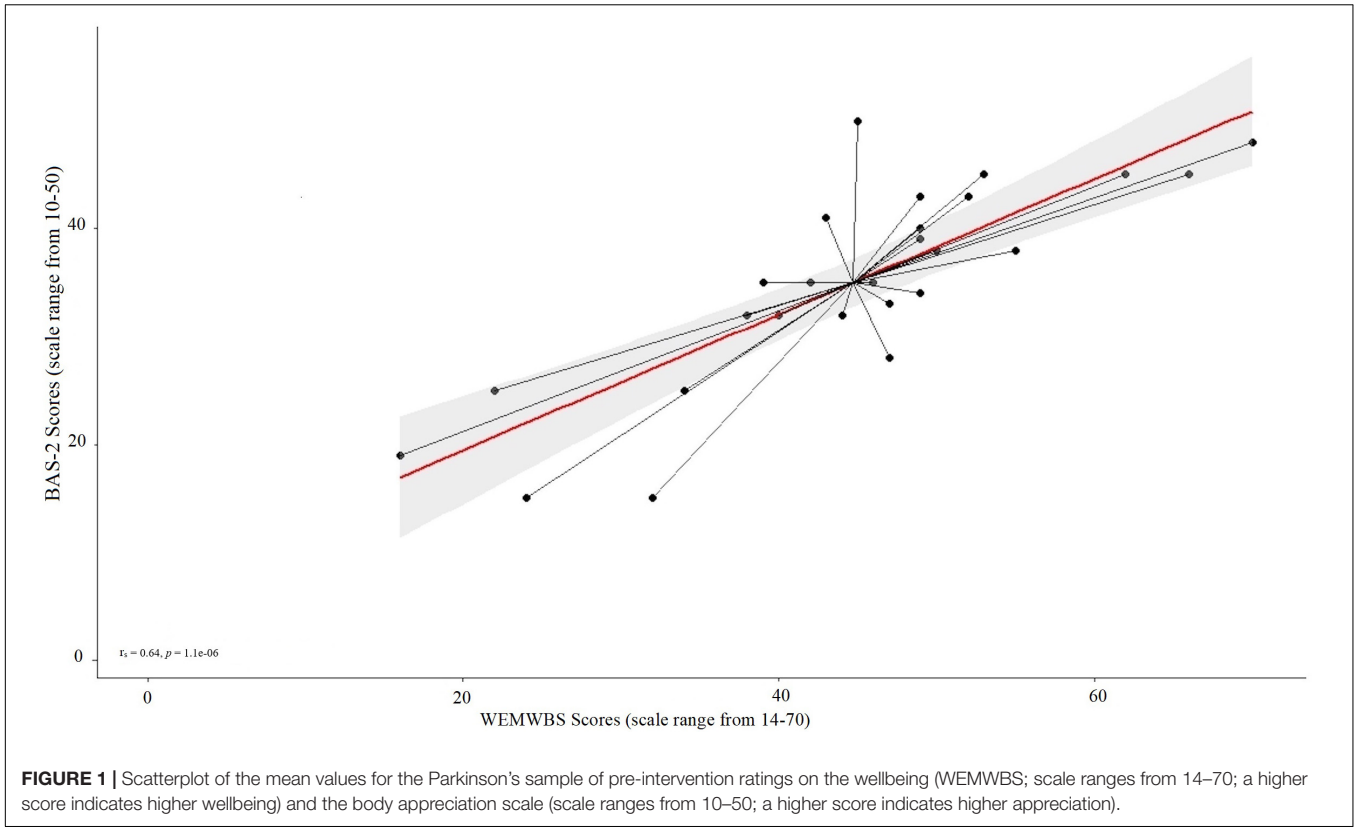

FIGURE 1 | Scatterplot of the mean values for the Parkinson's sample of pre-intervention ratings on the wellbeing (WEMWBS; scale ranges from 14–70; a higher score indicates higher wellbeing) and the body appreciation scale (scale ranges from 10–50; a higher score indicates higher appreciation).

A main effect of group was found on well-being scores (Z = -1.96, p = 0.05, r = 0.31), with controls showing higher levels of well-being in comparison with PwP (PwP: *median* = 48.00, *interquartile range* = 9.75, M = 46.3, SD = 12.29; controls: *median* = 55.00, *interquartile range* = 14.50, M = 54.42, SD = 8.81; **Figure 3**).

However, the interaction between group and time was not significant (Z = -1.46, p = 0.14, r = 0.23).

Body Appreciation

To investigate whether participation in a dance class shapes body appreciation (measured using the BAS-2) in PwP compared with age-matched controls, Mann-Whitney U-tests were used to analyze between-subjects factor of a 2 (groups: PwP vs. control) × 2 (time: pre vs. post) mixed design. Within-subjects factors were analyzed via Wilcoxon signed-rank tests. A significant main effect of time (Z = -2.62, p = 0.01 r = 0.42) was found, with post-dance scores higher than pre-dance ones (pre: *median* = 37.00, *interquartile range* = 12.00, M = 35.56, SD = 8.63; post: *median* = 38.00, *interquartile range* = 12.50, M = 38.1, SD = 11.55; **Figure 4**). No main effect of group was found (Z = -0.94, p = 0.35, r = 0.15).

A significant two-way interaction was revealed between group and time (Z = -2.05, p = 0.04, r = 0.33). Further investigation via Bonferroni-corrected *post hoc* tests indicated that the control group's body appreciation scores increased more following dance intervention than PwP (controls: Z = -2.72, p = 0.007, r = 0.53; pre: *median* = 38.00, *interquartile range* = 13.00, M = 36.69, SD = 6.79; post: *median* = 40.00, *interquartile range* = 9.00, M = 42.69, SD = 12.5; PwP: Z = -1.07, p = 0.28, r = 0.21; pre: *median* = 35.00, *interquartile range* = 10.50, M = 35, SD = 9.5; post: *median* = 37.50, *interquartile range* = 14.00, M = 35.81, SD = 10.55).

QUALITATIVE STUDY

Methods

Participants

Participants were recruited from a dance class at the University of Hertfordshire, local activity groups, and via word of mouth by those who attended the groups. To be eligible for the study, participants were required to have a diagnosis of Parkinson's and no other diagnosed movement or neurological condition. Initially, 10 PwP were recruited for the interviews, although only 8 were finally able to participate. All participants were taking dopaminergic medication, and none reported any changes to their medication in the weeks prior to participating in the study.

The participants included four individuals who regularly attended dance classes (range = 4–7 years) at the University of Hertfordshire (three female, age range = 46–70, M = 60.75, SD = 12.42). Two PwP had participated in the quantitative part of the study. The severity of Parkinson's ranged from mild to severe (Hoehn and Yahr stage 1–4) and years living with the condition ranged from 5 to 17. In addition to dance, the participants took part in a range of other activities, such as seated yoga, cycling, and an aerobic exercise group. The reported time spent being physically active ranged from 1–2 h to 3–4 h per week.

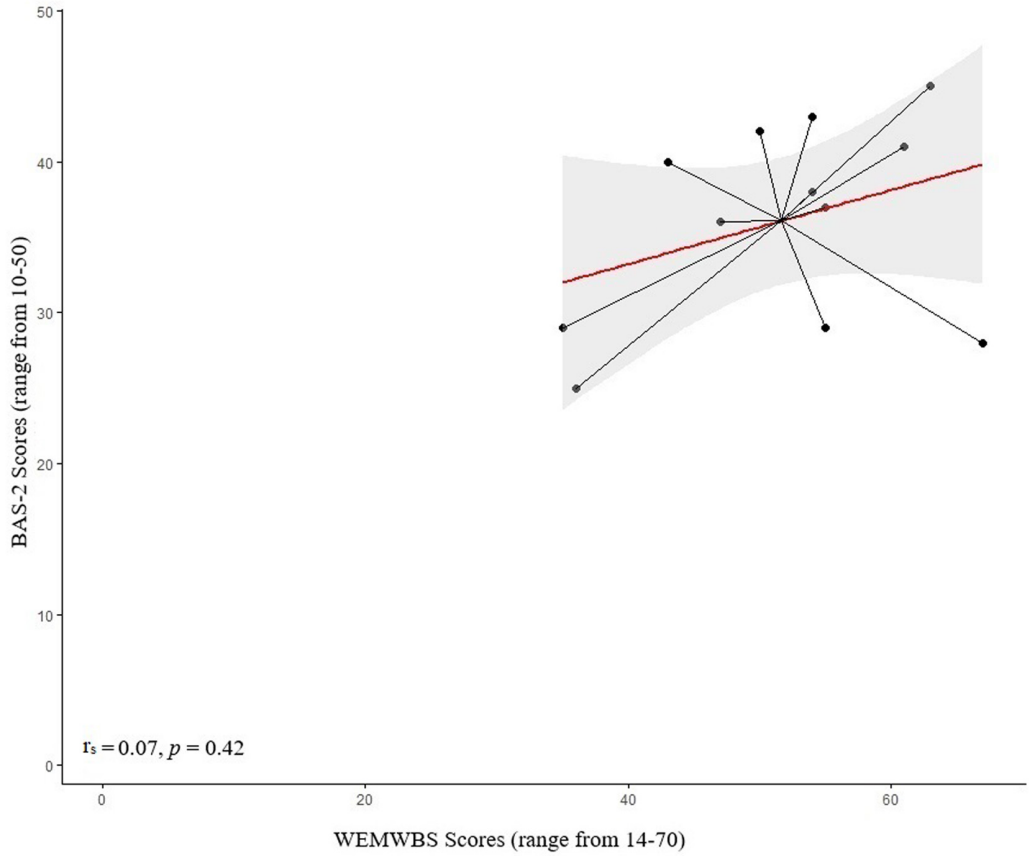

FIGURE 2 | Scatterplot of the mean values for the control sample of pre-intervention ratings on the wellbeing (WEMWBS; scale ranges from 14–70; a higher score indicates higher wellbeing) and the body appreciation scale (scale ranges from 10–50; a higher score indicates higher appreciation).

The remaining four participants (three male, age range = 50–78, $M = 67.75$, $SD = 12.42$) did not participate in the dance classes but engaged in other types of physical activity including golf, leisure walking, and aqua aerobics. The severity of Parkinson's ranged from mild to moderate (Hoehn and Yahr stage 1–3) and years living with the condition ranged from 2 to 14. The reported time spent being physically active ranged from 2–3 h to more than 4 h per week. This study was approved by an institutional Ethics Committee and all participants gave informed, written consent.

Interview Structure

Semi-structured interview schedules (see **Supplementary Material**) were developed based on reviewed research and sought to explore how participants felt about their bodies both physically and psychologically in the past and at present. Feelings toward the body may include thoughts about physical functionality, attitudes (Clancy, 2010), and appearance (Pila et al., 2016). The semi-structured approach was chosen because it allows for open-ended questions to gather data grounded in the experience of the participant while also incorporating questions guided by constructs of interest to the researchers (Galletta, 2013).

Interview Setting

Individual interviews were conducted with PwP who regularly participated in Dance for Parkinson's classes at the University of Hertfordshire and with PwP who exercised but did not dance. All interviews were conducted by a psychology student at the university campus. Two of the participants who did not dance requested that their partners be present during the interview; however, neither partner contributed to the interview. Each interview lasted approximately 30–45 min and was audiorecorded using an Olympus digital voice recorder.

Analysis

In order to maintain the full richness of the data and conversational nature of the interview, naturalized transcription was used (Oliver et al., 2005). In line with Braun and Clarke's (2006) analysis recommendations in the instance of an *a priori* research focus (the effect of dance on bodily awareness in the current study), the authors used theoretical thematic analysis to analyze the transcripts. The transcripts were subsequently coded independently by two psychologists (RH and DR) according to *a priori* hierarchical themes. This allows for the further development of previous research insights pertinent to the topic (Gamarra et al., 2009; Ghielen et al., 2017; Beerenbrock et al.,

2020). The five *a priori* codes were (1) dance and the body (Houston and McGill, 2013; Soundy et al., 2014; Koch et al., 2016; Beerenbrock et al., 2020), (2) body functionality, (3) body frustration, (4) pre/post-diagnosis changes in bodily awareness (Taleporos and McCabe, 2002), and (5) body dissatisfaction (Gamarra et al., 2009). Where utterances supported *a priori* codes, they were classified accordingly. Where text offered novel insight, new labels and descriptions were assigned. These were later discussed by the two researchers until agreement was reached regarding the nature of their contribution to the extant literature. In order to contextualize the utterance, participants who attended a dance class were titled "PD-dancer" (PwP who danced). Participants who did not dance were titled "PD-no-dance" (PwP who did not dance but took part in other movement-based activities).

Findings

The title and description of the *a priori* codes were reviewed and retitled, and subthemes were added to reflect the data in this study (**Supplementary Table 2**). Thematic mind maps were then used to explore the relationships between the *a priori* and novel codes (see **Supplementary Figures 6**, **7**). The respective findings are described in order of contribution to the following research questions: (1) How do participants perceive their body in relation to their condition in general? (2) How is awareness of the body experienced by participants who dance and those who take part in other movement-based activities? (3) Which aspects of dancing are related to participants perceptions of their body? Selected themes and subthemes that elaborate on the quantitative findings of the current study and add insight to the results of the reviewed literature are considered below (full results in **Supplementary Materials**).

Theme 1 Parkinson's and body perception

In relation to PwP general perspectives about their condition and its impact on their ability to engage in the activities they would like to do, two subthemes were identified (physical limitations and Parkinson's symptoms). First, the struggles and frustration both groups of participants felt with their own bodies due to their condition were evident, as illustrated in the following comments:

"Like getting dressed I get frustrated. You think I must be able to find the armholes"

(PD4-Dancer)

"It makes me tired. It makes me frustrated because I can't do things as well as I used to"

(PD1-Dancer)

"If it wasn't for the Parkinson's I'd be doing all sorts of things." (PD5-No-dance)

The second subtheme (Parkinson's symptoms) captures the uncertainty associated with symptom variability and the perception of lack of control being problematic, as the following examples convey:

"I guess the frustrating thing about Parkinson's is that you can do something one minute and then half an hour later you can't do it, and there's no explanation why you can't do it some days" (PD3-Dancer)

"I get very annoyed with my limbs at times. When I wash my hair in this kind of motion this hand often just stops, and I'm doing this and I'm thinking well, you know – do something!"

(PD7-No-dance)

However, some participants expressed being able to move past the frustration to a place of acceptance even with the unpredictable nature of their symptoms:

"I can accept that I have them [symptoms of Parkinson's] because I control them – because I have some agency over them. Whereas in the past, it was more difficult to accept and as they felt like they were beginning to control me" (PD7-No-dance).

Outside of movement-based activities, navigating everyday activities with Parkinson's highlighted how other peoples' lack of understanding regarding Parkinson's were at odds with participants own perception of their body's functionality.

Theme 2 Perceptions of the functionality of their own body

The second theme compares the impact of different perspectives and types of activities on body perception in relation to Parkinson's. Three contributing factors were found: (a) independence, (b) comparing to others, and (c) other people's understanding of PD.

The issue of independence observes the way in which both no-dance and dance participants perceived their body as an important factor in maintaining their autonomy.

Regardless of whether they danced or not, it was notable that participants described what their body enabled them to do rather than focusing on aesthetics, as the following extract demonstrates:

"I can still ride a two-wheel bicycle. So, my balance isn't too bad." (PD4-Dancer)

Participant PD7-No-dance listed several activities that they were able to do without needing help and expressed gratitude for the independence a functioning body gave them:

"I can walk around people and I can go through doorways without freezing. . .I can dance and I can [DJ] mix. I can make food. I am fine on self-care."

"I am grateful that I have a body that transports me where I want to go under my own steam."

The subtheme of comparing to others captures participants' perceptions of their body as an important factor in challenging their own view of Parkinson's. Comparisons between themselves and other PwP who attended the same activities were also a way to monitor their status; a litmus test for the progression of their own symptoms as the following excerpts demonstrate:

"*When you are with other people you realize their problems are greater than yours or otherwise if I think that's probably one guy*

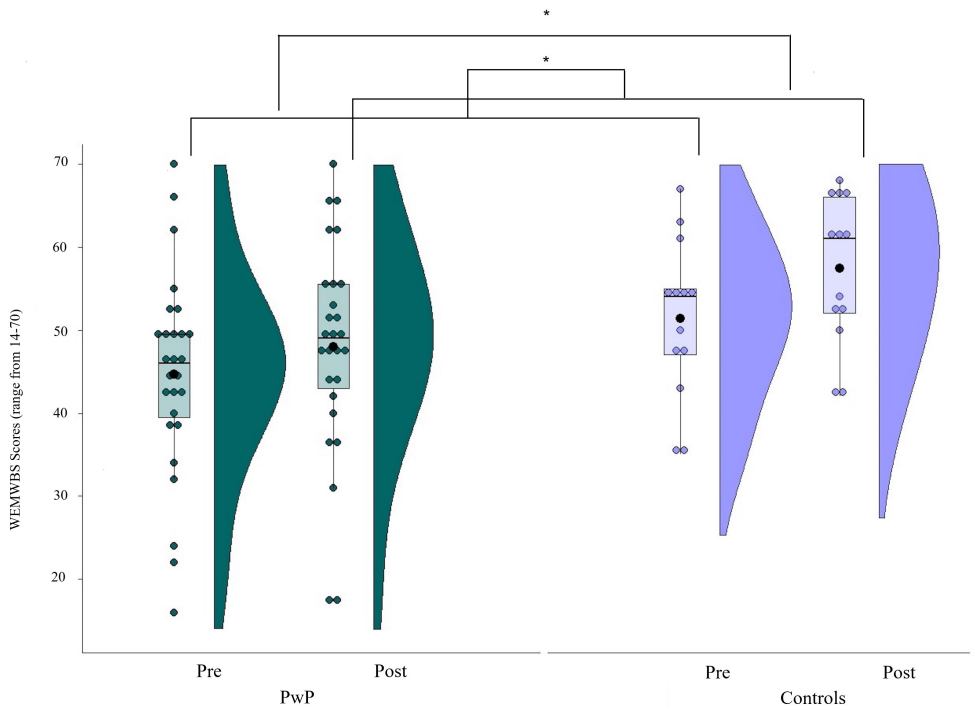

FIGURE 3 | Mean values of the wellbeing (WEMWBS). Scores on a scale from 14 (lower wellbeing) to 70 (higher wellbeing); Solid line = median; Black dot = mean; Whiskers: upper whisker = min(max(x), Q_3 + 1.5*IQR); lower whisker = max(min(x), Q_1−1.5*IQR.

that's better off than me. I mean, he's been diagnosed 3 years I've been diagnosed five so that's the difference, but you've got to take in what's going on around you." (PD8-No-dance)

"Being involved with other people that have Parkinson's who are far worse than I am makes me realize how lucky I am at the moment to be just at this stage after 6 years."(PD1-Dancer)

Theme 3 The impact of a dance class on bodily awareness

This theme captures PwP perceptions of dance classes as a space for bringing awareness to their ability to participate in the class (subtheme: negative awareness) and in which their preconceptions of Parkinson's were challenged (subtheme: comparison to others with PD).

In the dance class, comparisons with able-bodied student volunteers were experienced in a positive way as this exemplar shows:

"*The mix of able-bodied people and disabled people together is good and I quite like it when people like you [the researcher] can't do things.*" *(PD4-Dancer)*

Participating in dance drew attention in a negative way to changes in movement quality, as indicated below:

"*It reminds me that I'm getting old and I can't move as freely as I did." (PD1-Dancer)*

For some participants, this awareness meant coming to terms with uncontrollable changes associated with Parkinson's, as the following quote highlights:

"*I think if you do exercise you think that's going to fix me but then you realize that the fact that you can't do it isn't because you're not fit enough. . .[but] because it's coming from your brain.*" (*PD3-Dancer*)

Participants' uncertainty of their own future with Parkinson's contributed to an initial apprehension about attending a dance class:

"I would say that it was the only thing that really worried me before I came, that I was going to see a path mapped out in front of me." (PD2-Dancer)

However, the same participant described a sense of enablement from seeing others with PD successfully take part in the dance classes despite initial worries that they would experience the opposite effect:

"People that have been diagnosed and living with it way longer than me, are still getting up and dancing. In fact, it's amazingly encouraging, and it keeps you going really."

(PD2-Dancer)

Motivation was also gained from the opportunity to track the progression of their condition against other PwP at the class:

"*It's a kind of indicator of where you were at, where you are now, and kind of keep going.*" *(PD4-Dancer)*

Overall, the social aspect of attending specialized dance classes were described as very positive experiences.

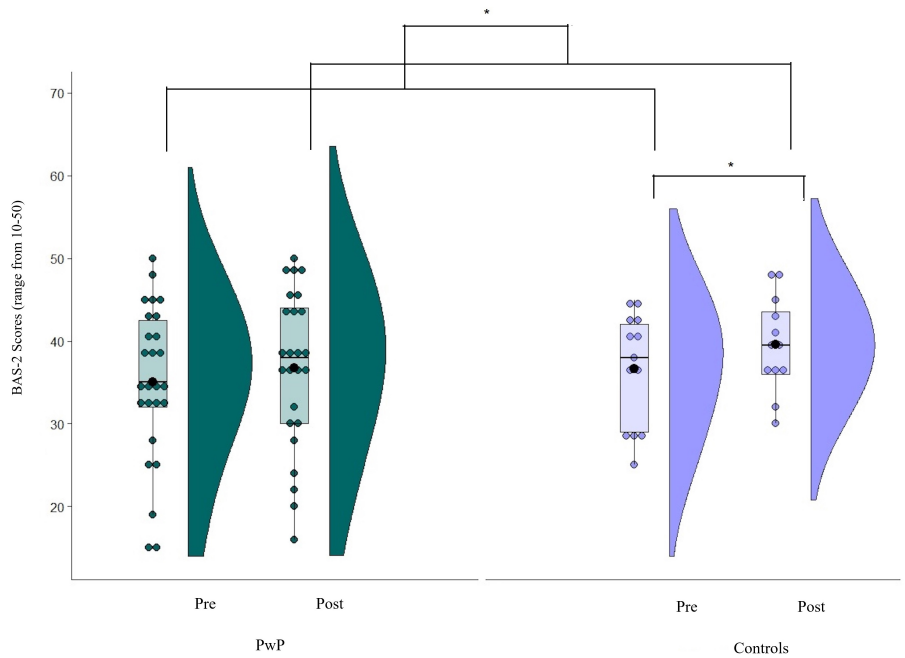

FIGURE 4 | Mean values of the body appreciation (BAS-2). Scores on a scale from 10 (lower body appreciation) to 50 (greater body appreciation); Solid line = median; Black dot = mean; Whiskers: upper whisker = min(max(x), Q_3 + 1.5*IQR); lower whisker = max(min(x), Q_1−1.5*IQR.

DISCUSSION

The current study explored the relationship between perceived well-being and body appreciation, as well as the effect of dance classes on both constructs in PwP and age-matched controls. We found that perceived well-being and body appreciation (pre-intervention) were significantly positively correlated in PwP but not in the control sample. Previous studies have shown a positive relationship between body appreciation and specific dimensions of perceived well-being in healthy individuals (e.g., self-esteem, Avalos et al., 2005; Swami et al., 2009 and self-compassion, Homan and Tylka, 2015; Marta-Simões and Ferreira, 2019). However, in the current study, the association between these constructs was specific to PwP rather than controls. This finding should be interpreted with caution given the small size of the control sample, and further research is needed to shed light on these conflicting results. To the best of our knowledge, no study to date has investigated the relationship between well-being and body appreciation in the context of PwP. Further research is needed to examine the potential clinical implications of body-focused interventions, such as dance, for PwP, taking into consideration the relationship between well-being and body appreciation.

The second aim of the current study was to investigate whether participation in a dance class would lead to improvements in perceived well-being and body appreciation. Findings suggest that taking part in dance classes improved well-being for PwP and controls. Previous studies have shown an effect of dance on mood (Lewis et al., 2016) and quality of life (Heiberger et al., 2011; Westheimer et al., 2015; Aguiar et al., 2016), even following a brief intervention (4 weeks; Blandy et al., 2015; singular workshop; Koch et al., 2016). The current findings confirm that dance, even as a singular session, may be a useful therapeutic intervention that is psychologically beneficial for all attendees. However, the improvement in well-being post-dance class was slightly more pronounced for controls compared to PwP. Few studies include a control group that undertake the same dance intervention as PwP when investigating the effect of dance on well-being. Therefore, future studies should explore the extent to which dance interventions may benefit perceived well-being for PwP relative to people without PD, such as the partners and caregivers of PwP.

Previous research with healthy individuals has shown regular dance participation to be associated with higher body appreciation (ballet and contemporary dance; Swami and Harris, 2012, street dance; Swami and Tovée, 2009) and enhanced body image (Burgess et al., 2006). In line with these studies, we found a greater increase in body appreciation scores for the age-matched control group following participation in a dance class, supporting the suggestion that dance may foster body enhancement in non-experts (Langdon and Petracca, 2010). However, no specific improvement was found in body appreciation for PwP. To the best of our knowledge, there are no other studies to date that have investigated the impact that dance interventions may have on perceived body appreciation for PwP. The lack of findings on body appreciation as reported in the current study may be due to a heightened bodily awareness following dance classes. It may be speculated that such an increase in bodily awareness may not necessarily translate into appreciation but may lead to insights into PwP's physical limitations. Preliminary qualitative findings

from the current study suggest that participants who danced experienced bodily awareness as a change in movement quality and an understanding that Parkinson's imposes limitations on their ability to participate in the class. Future studies could incorporate body compassion practices into interventions with PwP to help them transition from bodily awareness to body appreciation.

Qualitative findings of the present study suggest that participating in dance classes may draw negative awareness to the body in PwP, particularly in terms of functional capabilities and limitations, and that PwP tended to critically compare themselves with others. Therefore, dance interventions for PwP aiming to foster improvements in perceived well-being may benefit from incorporating elements that encourage body appreciation. Increased bodily awareness, including movement deficits and remaining mobility, have recently been reported by Beerenbrock et al. (2020) who explored the impact of tango classes on the body experience of PwP. Beerenbrock et al. (2020) suggest that body-related concerns are likely to be compounded by symptoms of the condition, such as impairments in balance and gait, which can lead to lower self-confidence in one's ability to carry out activities of daily living and a fear of falling (Palakurthi and Burugupally, 2019). The lack of control that PwP perceive to have over their body, as described by both the dance and no-dance participants in the qualitative study, alongside higher body-related concerns, may result in PwP becoming detached from their bodies (Beerenbrock et al., 2020). Hence, dance may lead to heightened bodily awareness, which may, in turn, negatively impact body appreciation following dance interventions. Future studies should investigate the extent to which increased bodily awareness may lead to higher or lower levels of body appreciation in PwP.

The findings of this current exploratory study should be considered in light of some contextual limitations. With regard to recruitment, it should be noted that some participants regularly attend dance classes, and this was not a novel experience, a factor that may have led to a positive bias toward dance. Due to the lack of systematic control over participants' dance experience, it is not known whether the effect of dance on well-being and body appreciation can be generalized to novice dancers. The effect found in the current study may be limited to participants who had more experience of dancing. Future studies should consider controlling for participant experience with dancing. The results of the quantitative study highlight that participation in a singular hour-long dance class may result in immediate benefit in terms of well-being and body appreciation. However, from these results, it is not known whether this effect is consistent or remains over time. Future studies should include a long-term follow-up to address this. The quantitative study included a limited control group that took part in the same dance class as the PwP. The inclusion of another group undertaking a different type of exercise would shed light on whether changes in well-being and body appreciation are specific to dance or exercise more generally. Given that the control group predominantly included of PwP partners, it could be suggested that the well-being of controls was improved by proxy due to perceived improvement in the well-being of their partner with PD. While there is some supporting qualitative evidence for this argument (Prado et al., 2020), we believe that this would reflect an important finding in itself, as it indicates the value in attending dance for both groups. In particular, the findings of Prado et al. (2020) also indicated that caregivers may attend dance classes, as there is not enough time for them to do something else on their own. Therefore, joint interventions such as dance may provide an important opportunity that might not otherwise be afforded for some caregivers to benefit, psychologically, even if this is through the improved well-being of their partner. The control group was also not of equal size to the Parkinson's group. The discrepancy in sample size between the Parkinson's and control group is in line with several other studies that have investigated the physical and psychological benefits of dance for PwP (Romenets et al., 2015; Lewis et al., 2016) and reflects a genuine imbalance between those who attend classes with partners and those who do not (Prado et al., 2020). The quantitative study did not include a no-dance PD control group. Future studies could consider including people with greater disease progression and more severe symptoms in order to understand how the nature of this degenerative condition impacts on individuals' bodily concerns in relation to dance. Finally, the qualitative study was limited to interviews with eight PwP; therefore, the results should be treated as preliminary. Future studies should consider including a larger sample size when exploring the effect of dance on bodily awareness in Parkinson's.

Overall, a positive relationship between dancing and well-being was observed in this study. However, no effect of dance was found on body appreciation in PwP, and qualitative data suggested that the participants' focus was on potential physical limitations, rather than body appreciation, which was heightened by the comparison with other classes' attendees. Taken together, these findings suggest that dance interventions aiming to foster well-being in PwP should take into consideration the repercussions on perceived body appreciation, by adopting a more bodily positive attitude throughout the intervention and explicitly tackling its potential impact on PwP.

DATA AVAILABILITY STATEMENT

The datasets generated for this study are available on request to the corresponding author.

ETHICS STATEMENT

The studies involving human participants were reviewed and approved by the Health, Sciences, Engineering and Technology Ethics Committee with Delegated Authority at the University of Hertfordshire. The participants provided their written informed consent to participate in this study.

AUTHOR CONTRIBUTIONS

OE-G conducted the interviews, transcribed the qualitative data, and contributed to the manuscript. RH and SP collected the

quantitative data and supervised the study. Qualitative data analysis was undertaken by RH and DR and quantitative by SP and RH. MK aided the recruitment of participants from local dance classes and reviewed the manuscript. Revisions of the manuscript were undertaken by RH, SP, and DR. All authors contributed to the article and approved the submitted version.

FUNDING

RH undertook this research while funded by the Ph.D. studentship from the University of Hertfordshire.

ACKNOWLEDGMENTS

We would like to thank the people with Parkinson's who gave their time to participate in the current study. We would also like to acknowledge Kezia Jolly for allowing us to recruit participants from her dance classes.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fpsyg.2020.531567/full#supplementary-material

REFERENCES

Aguiar, L. P. C., da Rocha, P. A., and Morris, M. (2016). Therapeutic dancing for Parkinson's disease. *Int. J. Gerontol.* 10, 64–70. doi: 10.1016/j.ijge.2016.02.002

Allen, M., Poggiali, D., Whitaker, K., Marshall, T. R., and Kievit, R. A. (2019). Raincloud plots: a multi-platform tool for robust data visualization. *Wellcome Open Res.* 4:63.

Avalos, L., Tylka, T. L., and Wood-Barcalow, N. (2005). The body appreciation scale: development and psychometric evaluation. *Body Image* 2, 285–297. doi: 10.1016/j.bodyim.2005.06.002

Beerenbrock, Y., Meyer, L., Böhme, J., Herrlich, S., Mews, S., Berger, B., et al. (2020). Perceived effects of tango argentino on body experience in persons with Parkinson's disease (PD)—a qualitative study with affected persons and their partners. *Complement. Ther. Med.* 48:102221. doi: 10.1016/j.ctim.2019.102221

Blandy, L. M., Beevers, W. A., Fitzmaurice, K., and Morris, M. E. (2015). Therapeutic argentine tango dancing for people with mild Parkinson's disease: a feasibility study. *Front. Neurol.* 6:122. doi: 10.3389/fneur.2015.00122

Braun, V., and Clarke, V. (2006). Using thematic analysis in psychology. *Qualitat. Res. Psychol.* 3, 77–101. doi: 10.1191/1478088706qp063oa

Burgess, G., Grogan, S., and Burwitz, L. (2006). Effects of a 6-week aerobic dance intervention on body image and physical self-perceptions in adolescent girls. *Body Image* 3, 57–66.

Chaudhuri, K. R., Odin, P., Antonini, A., and Martinez-Martin, P. (2011). Parkinson's disease: the non-motor issues. *Parkins. Related Disord.* 17, 717–723. doi: 10.1016/j.parkreldis.2011.02.018

Christopher, L., Koshimori, Y., Lang, A. E., Criaud, M., and Strafella, A. P. (2014). Uncovering the role of the insula in non-motor symptoms of Parkinson's disease. *Brain* 137, 2143–2154.

Clancy, S. E. (2010). *The Effects of Yoga on Body Dissatisfaction, Self-objectification, and Mindfulness of the Body in College Women*. Ph.D. Dissertation, Washington State University, Pullman, Washington, DC.

Courtois, I., Cools, F., and Calsius, J. (2015). Effectiveness of body awareness interventions in fibromyalgia and chronic fatigue syndrome: a systematic review and meta-analysis. *J. Bodywork Movement Ther.* 19, 35–56. doi: 10.1016/j.jbmt.2014.04.003

Craig, A. D. (2002). How do you feel? Interoception: the sense of the physiological condition of the body. *Nat. Rev. Neurosci.* 3, 655–666. doi: 10.1038/nrn894

Dorsey, E. R., Elbaz, A., Nichols, E., Abd-Allah, F., Abdelalim, A., Adsuar, J. C., et al. (2018). Global, regional, and national burden of Parkinson's disease, 1990–2016: a systematic analysis for the Global Burden of Disease Study 2016. *Lancet Neurol.* 17, 939–953. doi: 10.1016/S1474-4422(18)30295-3

Duncan, G. W., Khoo, T. K., Yarnall, A. J., O'Brien, J. T., Coleman, S. Y., Brooks, D. J., et al. (2014). Health-related quality of life in early Parkinson's disease: the impact of nonmotor symptoms: quality of life in early Parkinson's disease. *Movement Disord.* 29, 195–202. doi: 10.1002/mds.25664

Galletta, A. (2013). *Mastering the Semi-Structured Interview and Beyond: From Research Design to Analysis and Publication*. New York, NY: New York University Press.

Gamarra, A. H. E., Molski, C. S., Hilbig, A., Valentini, C. K., Striebel, V. L. W., Rieder, C. R., et al. (2009). Evaluation of body image and self-concept and their correlation with depressive symptoms in Parkinson's disease. *Arquivos Neuro Psiquiatria* 67, 585–590.

Ghielen, I., van Wegen, E. E., Rutten, S., de Goede, C. J., Houniet-de Gier, M., Collette, E. H., et al. (2017). Body awareness training in the treatment of wearing-off related anxiety in patients with Parkinson's disease: results from a pilot randomized controlled trial. *J. Psycho. Res.* 103, 1–8.

Hackney, M. E., and Earhart, G. M. (2009). Health-related quality of life and alternative forms of exercise in Parkinson disease. *Parkins. Related Disord.* 15, 644–648. doi: 10.1016/j.parkreldis.2009.03.003

Heiberger, L., Maurer, C., Amtage, F., Mendez-Balbuena, I., Schulte-Mönting, J., Hepp-Reymond, M.-C., et al. (2011). Impact of a weekly dance class on the functional mobility and on the quality of life of individuals with Parkinson's disease. *Front. Aging Neurosci.* 3:1–15. doi: 10.3389/fnagi.2011.00014

Holmes, W. M., and Hackney, M. E. (2017). Adapted tango for adults with Parkinson's disease: a qualitative study. *Adapted Phys. Activity Q.* 34, 256–275. doi: 10.1123/apaq.2015-0113

Homan, K. J., and Tylka, T. L. (2015). Self-compassion moderates body comparison and appearance self-worth's inverse relationships with body appreciation. *Body Image* 15, 1–7. doi: 10.1016/j.bodyim.2015.04.007

Houston, S., and McGill, A. (2013). A mixed-methods study into ballet for people living with Parkinson's. *Arts Health* 5, 103–119. doi: 10.1080/17533015.2012.745580

Jenkinson, C., Peto, V., Fitzpatrick, R., Greenhall, R., and Hyman, N. (1995). Self-reported functioning and well-being in patients with Parkinson's disease: comparison of the short-form health survey (SF-36) and the Parkinson's Disease Questionnaire (PDQ-39). *Age and Ageing* 24, 505–509.

Kalia, L. V., and Lang, A. E. (2015). Parkinson's disease. *Lancet* 386, 896–912. doi: 10.1016/S0140-6736(14)61393-3

Kim, S., and Kim, J. (2007). Mood after various brief exercise and sport modes: aerobics, hip-hop dancing, ice skating, and body conditioning. *Percept. Motor Skills* 104, 1265–1270.

Koch, S. C., Mergheim, K., Raeke, J., Machado, C. B., Riegner, E., Nolden, J., et al. (2016). The embodied self in Parkinson's disease: feasibility of a single tango intervention for assessing changes in psychological health outcomes and aesthetic experience. *Front. Neurosci.* 10:287. doi: 10.3389/fnins.2016.00287

Koller, W. C. (1984). Sensory symptoms in Parkinson's disease. *Neurology* 34, 957–957.

Kolnes, L.-J. (2017). Exercise and physical therapy help restore body and self in clients with severe anorexia nervosa. *J. Bodywork Movement Ther.* 21, 481–494. doi: 10.1016/j.jbmt.2016.09.005

Langdon, S. W., and Petracca, G. (2010). Tiny dancer: body image and dancer identity in female modern dancers. *Body Image* 7, 360–363. doi: 10.1016/j.bodyim.2010.06.005

Lewis, C., Annett, L. E., Davenport, S., Hall, A. A., and Lovatt, P. (2016). Mood changes following social dance sessions in people with Parkinson's disease. *J. Health Psychol.* 21, 483–492. doi: 10.1177/1359105314529681

Marta-Simões, J., and Ferreira, C. (2019). Turning self-compassionate engagement into action: impact on body appreciation among Portuguese women. *J. Contextual Behav. Sci.* 11, 21–26. doi: 10.1016/j.jcbs.2018.11.004

Mehling, W. E., Chesney, M. A., Metzler, T. J., Goldstein, L. A., Maguen, S., Geronimo, C., et al. (2018). A 12-week integrative exercise program improves self-reported mindfulness and interoceptive awareness in war veterans with posttraumatic stress symptoms. *J. Clin. Psychol.* 74, 554–565. doi: 10.1002/jclp.22549

Nicoletti, A., Mostile, G., Stocchi, F., Abbruzzese, G., Ceravolo, R., Cortelli, P., et al. (2017). Factors influencing psychological well-being in patients with Parkinson's disease. *PLoS One* 12:e0189682. doi: 10.1371/journal.pone.0189682

Nimwegen, M., Speelman, A. D., Hofman-van Rossum, E. J. M., Overeem, S., Deeg, D. J. H., Borm, G. F., et al. (2011). Physical inactivity in Parkinson's disease. *J. Neurol.* 258, 2214–2221. doi: 10.1007/s00415-011-6097-7

Oliver, D. G., Serovich, J. M., and Mason, T. L. (2005). Constraints and opportunities with interview transcription: towards reflection in qualitative research. *Soc. Forces* 84, 1273–1289.

Palakurthi, B., and Burugupally, S. P. (2019). Postural Instability in Parkinson's disease: a review. *Brain Sci.* 9:239. doi: 10.3390/brainsci9090239

Pfeiffer, R. F. (2016). Non-motor symptoms in Parkinson's disease. *Parkins. Related Disord.* 22, S119–S122. doi: 10.1016/j.parkreldis.2015.09.004

Pila, E., Brunet, J., Crocker, P. R. E., Kowalski, K. C., and Sabiston, C. M. (2016). Intrapersonal characteristics of body-related guilt, shame, pride, and envy in Canadian adults. *Body Image* 16, 100–106. doi: 10.1016/j.bodyim.2016.01.001

Prado, L., Hadley, R., and Rose, D. (2020). Taking time: a mixed methods study of Parkinson's disease caregiver participation in activities in relation to their wellbeing. *Parkinson's Dis.* 2020:7370810. doi: 10.1155/2020/7370810

Pringsheim, T., Jette, N., Frolkis, A., and Steeves, T. D. L. (2014). The prevalence of Parkinson's disease: a systematic review and meta−analysis. *Movement Disord.* 29, 1583–1590. doi: 10.1002/mds.25945

Reynolds, G. O., Otto, M. W., Ellis, T. D., and Cronin-Golomb, A. (2016). The therapeutic potential of exercise to improve mood, cognition, and sleep in Parkinson's disease: exercise and nonmotor symptoms of PD. *Movement Disord.* 31, 23–38. doi: 10.1002/mds.26484

Romenets, S. R., Anang, J., Fereshtehnejad, S.-M., Pelletier, A., and Postuma, R. (2015). Tango for treatment of motor and non-motor manifestations in Parkinson's disease: a randomized control study. *Complement. Ther. Med.* 23, 175–184. doi: 10.1016/j.ctim.2015.01.015

Schapira, A. H. V., Chaudhuri, K. R., and Jenner, P. (2017). Non-motor features of Parkinson disease. *Nat. Rev. Neurosci.* 18, 435–450. doi: 10.1038/nrn.2017.62

Shanahan, J., Morris, M. E., Bhriain, O. N., Saunders, J., and Clifford, A. M. (2015). Dance for people with parkinson disease: what is the evidence telling us? *Arch. Phys. Med. Rehabil.* 96, 141–153. doi: 10.1016/j.apmr.2014.08.017

Soundy, A., Stubbs, B., and Roskell, C. (2014). The experience of Parkinson's disease: a systematic review and meta-ethnography. *Sci. World J.* 2014, 1–19. doi: 10.1155/2014/613592

Swami, V., Airs, N., Chouhan, B., Amparo Padilla, Leon, M., and Towell, T. (2009). Are there ethnic differences in positive body image among female British undergraduates? *Eur. Psychol.* 14, 288–296.

Swami, V., and Harris, A. S. (2012). Dancing toward positive body image? Examining body-related constructs with ballet and contemporary dancers at different levels. *Am. J. Dance Ther.* 34, 39–52. doi: 10.1007/s10465-012-9129-7

Swami, V., and Tovée, M. J. (2009). A comparison of actual-ideal weight discrepancy, body appreciation, and media influence between street-dancers and non-dancers. *Body Image* 6, 304–307. doi: 10.1016/j.bodyim.2009.07.006

Taleporos, G., and McCabe, M. P. (2002). Body image and physical disability—personal perspectives. *Social Sci. Med.* 54, 971–980. doi: 10.1016/S0277-9536(01)00069-7

Tennant, R., Hiller, L., Fishwick, R., Platt, S., Joseph, S., Weich, S., et al. (2007). The warwick-edinburgh mental well-being scale (WEMWBS): development and UK validation. *Health Quality Life Outcomes* 5:63. doi: 10.1186/1477-7525-5-63

Tylka, T. L., and Wood-Barcalow, N. L. (2015). The body appreciation scale-2: item refinement and psychometric evaluation. *Body Image* 12, 53–67. doi: 10.1016/j.bodyim.2014.09.006

Vankova, H., Holmerova, I., Machacova, K., Volicer, L., Veleta, P., and Celko, A. M. (2014). The effect of dance on depressive symptoms in nursing home residents. *J. Am. Med. Directors Assoc.* 15, 582–587.

Walter, O., and Yanko, S. (2018). New observations on the influence of dance on body image and development of eating disorders. *Res. Dance Educ.* 19, 240–251.

Westheimer, O., McRae, C., Henchcliffe, C., Fesharaki, A., Glazman, S., Ene, H., et al. (2015). Dance for PD: a preliminary investigation of effects on motor function and quality of life among persons with Parkinson's disease (PD). *J. Neural Transmission* 122, 1263–1270. doi: 10.1007/s00702-015-1380-x

Conflict of Interest: MK leads the dance classes that take place at the University of Hertfordshire.

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 in Psychology

HYPOTHESIS AND THEORY
published: 12 January 2021
doi: 10.3389/fpsyg.2020.531688

Into Your (S)Kin: Toward a Comprehensive Conception of Empathy

Tue Emil Öhler Søvsø[1,2] and Kirstin Burckhardt[3]*

[1] Department of Philosophy and Humanities, Institute of Greek and Latin Languages and Literatures, Freie Universität Berlin, Berlin, Germany, [2] Department of Philosophy, Humboldt Universität zu Berlin, Berlin, Germany, [3] Clinic for Psychosomatic Medicine and Psychotherapy, Städtisches Klinikum Görlitz, Görlitz, Germany

This paper argues for a comprehensive conception of empathy as comprising epistemic, affective, and motivational elements and introduces the ancient Stoic theory of attachment (Greek, *oikeiōsis*) as a model for describing the embodied, emotional response to others that we take to be distinctive of empathy. Our argument entails that in order to provide a suitable conceptual framework for the interdisciplinary study of empathy one must extend the scope of recent "simulationalist" and "enactivist" accounts of empathy in two important respects. First, against the enactivist assumption that human mindreading capacities primarily rely on an immediate, quasi-perceptual understanding of other's intentional states, we draw on Alfred Schutz' analysis of social understanding to argue that reflective types of understanding play a distinct, but equally fundamental role in empathic engagements. Second, we insist that empathy also involves an affective response toward the other and their situation (as the empathizer perceives this). We suggest analyzing this response in terms of the Stoic concepts of attachment, concern, and a fundamental type of prosocial motivation, that can best be described as an "extended partiality." By way of conclusion, we integrate the above concepts into a comprehensive conceptual framework for the study of empathy and briefly relate them to current debates about empathic perception and prosocial motivation. The result, we argue, is an account that stays neutral with regard to the exact nature of the processes involved in producing empathy and can therefore accommodate discussion across theoretical divides—e.g., those between enactivist, simulationalist, and so-called theory-theorist approaches.

Keywords: empathy, stoicism, embodied cognition, attachment, phenomenology, prosocial motivation, affective intentionality

OPEN ACCESS

Edited by:
Anna Ciaunica,
University of Porto, Portugal

Reviewed by:
Sara Dellantonio,
University of Trento, Italy
Joona Taipale,
University of Jyväskylä, Finland

****Correspondence:***
Tue Emil Öhler Søvsø
tue@zedat.fu-berlin.de

Specialty section:
This article was submitted to Theoretical and Philosophical Psychology, a section of the journal Frontiers in Psychology

Received: *04 February 2020*
Accepted: *09 December 2020*
Published: *12 January 2021*

Citation:
Søvsø TEÖ and Burckhardt K (2021) Into Your (S)Kin: Toward a Comprehensive Conception of Empathy. Front. Psychol. 11:531688. doi: 10.3389/fpsyg.2020.531688

INTRODUCTION

"Slipping into someone else's skin"—this bodily metaphor serves as an almost dictionary definition of empathy. In general, the body looms large when speaking about empathy and kindness. We "put ourselves in someone else's shoes" or "listen to our hearts" and the excessively empathic among us have "a bleeding heart." Nonetheless, the body occupies a somewhat un-easy place in modern discussions of empathy, which have tended to focus on just one part of the bodies involved, namely the brain of the empathizer.

With inspiration from phenomenology and interaction theory, recent accounts of empathy have pushed to transcend this focus on a more or less isolated mind, emphasizing the immediate, and "intercorporeal" nature of empathic interaction (see e.g., Fuchs and De Jaegher, 2009; Zahavi, 2011, 2014; Gallagher, 2012a; Ciaunica, 2017; Fuchs, 2017; for a good review, see Zahavi and Michael, 2018). These phenomenological and "enactivist" accounts construe empathy as a basic type of other-directed intentionality that enables us to directly perceive the mental states of others through our interaction with them. They thus form part of a broader research agenda known as 4E cognition committed to studying our cognitive processes as embodied, extended, enactive, and embedded phenomena (for a concise introduction, see Newen et al., 2018).

This conception of the mind as encompassing and in part co-constituted by the body and its environment carries great prospects for the study of interpersonal processes but arguably 4E analyses of empathy remain underdeveloped in crucial respects. While these phenomenologically inspired discussions of empathy have broadened the scope on empathy by conceiving of it as an embodied and enactive phenomenon, they have also tended to focus exclusively on its epistemic aspects neglecting the affective aspects at the very root of the word and broadly considered to be central to empathy. The aim of the present paper is to establish a conceptual framework that allows for a more comprehensive analysis of empathy as both an epistemic and an emotional phenomenon.

More specifically we propose to supplement the enactivist account of basic empathy with an analysis of the affective attitude and response involved in empathic engagements. We develop this analysis with inspiration from the ancient Stoics' account of affective intentionality, known as their theory of *oikeiōsis*. On this theory, the bodily processes, pointed to in enactivist accounts of empathy as enabling an immediate other-understanding, are instead seen as producing an attitude of attachment, concern, and extended partiality which provides the affective foundations of interpersonal relations. The Stoic analysis of embodied affectivity therefore nicely complements the traditional phenomenological analysis of empathy's epistemic aspects and provides the conceptual means for relating the enactivist account to the ongoing debates about empathy's role in motivating prosocial behavior.

In the section "Empathy: Cognitive, Emotional, Primary, and Extended" by briefly situating the enactivist approach within the broader debates about empathy and social understanding and introduce one of the most extensive accounts of enactive empathy, namely the one developed by the German philosopher and psychiatrist Thomas Fuchs. In the section "Social Understanding: Enactive and Reflective" we discuss the concept of bodily resonance, taken by Fuchs to underlie our empathic capacities, in more detail drawing on the early phenomenologist Alfred Schutz's (1967) analysis of social understanding and meaning to point out the limits of such "enactive" understanding and contrast it with explicit, "reflective" types of social understanding. In the subsequent section, we then introduce the Stoic theory of *oikeiōsis* and interpersonal relations and compare this account of affective intentionality with the one underlying Fuchs' account of empathy. Having thus delineated the epistemic and affective scope of empathy, in the section "Discussion: Toward a Comprehensive Conception of Empathy" we piece together our account of empathy specifying the individual processes we take it to involve and briefly relating it to existing, empirical debates about the attachment toward one's own body (so-called "Body Ownership") and the nature of the prosocial motivation produced by empathy.

EMPATHY: COGNITIVE, EMOTIONAL, PRIMARY, AND EXTENDED

Before zooming in on the many debates surrounding the concept of empathy, it may be useful to get a rough idea about what kind of thing(s) the term "empathy" is generally taken to refer to and how we shall use it. Batson (2009) has distinguished two questions that students of empathy have been interested in: "How can one know what another person is thinking and feeling?" and "What leads one person to respond with sensitivity and care to the suffering of another?" (Batson, 2009, p. 3). He identifies eight different psychological states that have all been called empathy and seen as relevant to answering one or both of the questions above. These states are often subsumed under the terms emotional and cognitive empathy, i.e., the ability to resonate with the emotions of others or "feel with them" and the appreciation of their emotions through purely cognitive means, also called perspective taking. Batson sees this complexity as a basic fact about the study of empathy that one needs to acknowledge and deal with as best one can.

Our strategy for dealing with this complexity is to distinguish between empathy as a complex emotional phenomenon and the various empathic processes that may go into producing it. This means that we will take a top-down approach to empathy analyzing it in its developed form as it appears relatively late in childhood and largely stay at the purely conceptual level of determining what epistemic, emotional, and motivational states empathy involves.

We tentatively define empathy as *a benevolent engagement with the affective states of others which provides us with a grasp of their state and produces an affective response within our body*. This engagement, we take it, plausibly involves a whole range of distinct, empathic processes, and states, which complement, interrelate, and inform each other in intricate ways, which we are really only beginning to grasp and describe scientifically (Batson, 2009; Zaki and Ochsner, 2012). The purpose of viewing these various processes as elements of empathy and not in isolation is to allow us to study their interaction and interrelations, not just their individual workings, on the assumption that there might be more to the whole than the sum of its parts (cf. Zaki and Ochsner, 2012).

Our definition of empathy is deliberately vague, and in the first instance, it only aims to provide a possibly neutral language for speaking about the individual aspects of empathy while allowing us to distinguish it from distinct, but closely related phenomena

such as emotional contagion, sharing, and mindreading[1]. To delineate the scope of our definition it may be useful to compare it to some recent, prominent attempts to define empathy.

Frédérique de Vignemont, along with shifting co-authors, has developed a highly systematic account of empathy and proposed a set of necessary and sufficient conditions that distinguish empathy from related phenomena (de Vignemont and Jacob, 2012; cf. de Vignemont and Singer, 2006). These are:

1. the affectivity condition: both the empathizer and the target of her empathy experiences an affective state
2. the interpersonal similarity condition: empathizer and target person must experience similar affective states
3. the causal path condition: the affective state of the empathizer must be caused by the affective state of the target person
4. the ascription condition: the empathizer must be aware of the target person's affective state and of the fact that this state is the cause of her own affective state
5. the caring condition: the empathizer must care about the target person's affective life

Our definition retains condition (i), (iii), and (v) by insisting that for an engagement to qualify as empathic it must be concerned with the affective state of someone else, it must produce an "affective response" in the empathizer, and it must be "benevolent[2]." Similar to de Vignemont, we thus distinguish empathy from standard mindreading (in the sense of merely registering the intentional states of others) and what is sometimes simply called sharing [where two people experience the same affective state but caused by the same intentional object (Scheler, 1954)].

Condition (iv) we only accept in a slightly weaker form. By speaking of a grasp rather than an explicit awareness of the other's state we thus wish to include the pre-reflective sensing and reacting to others' states, sometimes called "basic empathy" (Stueber, 2006; Fernandez and Zahavi, 2020), as instances of empathy along with the conscious ascriptions exclusively recognized by de Vignemont. This pre-reflective awareness of the other's state is still sufficient to distinguish empathy from emotional contagion, i.e., the unconscious adoption of others' affective states as witnessed e.g., in the spreading of mass panic (Scheler, 1954), which does not imply any awareness of others' states.

Condition (ii), the interpersonal similarity condition, is far more controversial and in effect limits the scope of empathy to cases of vicarious sharing. De Vignemont describes her account as "simulation-based" and commits herself to the view that empathy relies on the recognition of mental states through an internal simulation or enactment-imagination of that state, which is then ascribed to the other (de Vignemont and Jacob, 2012). This narrowly simulationalist conception of empathy is problematic since it arguably misconstrues the scope of empathic understanding. As Dan Zahavi argues: "To insist that the empathizer must have the same (kind of) state as the target, is to miss what is distinctive about empathy, namely the fact that it confronts you with the presence of an experience that you are not living through yourself" (Zahavi and Michael, 2018, p. 597).

By contrast, Zahavi has suggested a definition of empathy as "a distinctive form of other-directed intentionality, distinct from both self-awareness and ordinary object-intentionality, which allows foreign experiences to disclose themselves as foreign rather than as own" (Zahavi, 2014, p. 138). This definition of empathy as a distinct type of intentionality, in turn, restricts the scope empathy to its epistemic aspects (i.e., Batson's first question) and allows for empathy to appear in combination with any number of affective states and attitudes toward others. Indeed, Zahavi explicitly bites the bullet that on such an account "it is perfectly coherent to think that an expert torturer may rely on empathy in order to work out how best to push her victim's buttons" (Zahavi and Michael, 2018, p. 597).

This narrowly epistemic conception of empathy has a long tradition within philosophy—especially phenomenological philosophy where the concept originally evolved (Stueber, 2006; Zahavi, 2010)—and it creates a neat distinction between empathy and such clearly affective phenomena as sympathy or compassion. But it also blurs the line between empathy and mindreading more broadly and goes against the common intuition that empathy somehow implies a basic pro-attitude toward the other. Since this intuition is strong enough to have inspired an entire strand of research on empathy (i.e., the studies of Batson's second question), we believe that it should be reflected in the definition of the concept.

In contrast to Zahavi and following de Vignemont, we thus agree that empathy is a basically benevolent way of engaging with others where both empathizer and target person experience an affective state. Unlike de Vignemont, however, we do not assume that these states must be similar or that the state of the empathizer has to serve as her basis for understanding the target's state. It is sufficient that it is a response elicited by the engagement with that state (through simulation or otherwise).

In defining empathy as an engagement which both reveals the target's affective state *and* produces an affective response in the empathizer, we thus distinguish empathy from mindreading more broadly both in terms of its proper object and its effect. Empathy, on our view, is a type of mindreading that provides understanding of the affective states of others (widely construed as any state that involves the ascription of affective value, cf. Fuchs' concept of emotions discussed below, section "Social Understanding: Enactive, and Reflective"), whereas mindreading as such can deal with all types of mental states, including non-valenced beliefs. Also, empathy requires you to feel something for the person you empathize with, not just neutrally (or maliciously) registering their affective state (cf. Gallagher, 2012a who talks of empathy as involving a "primary and irreducible affective

[1]We shall use the term "mindreading" in the neutral sense of any capacity to register the intentional states of others, not as referring to a purely mentalistic understanding.

[2]On de Vignemont's account, condition (v) is introduced in order to account for the modulation of empathy by contextual and personal factors but it remains somewhat unclear what this condition involves apart from the fact that de Vignemont takes empathy to be incompatible with "an egotist stand" and likely to be blocked by an indifferent or negative attitude toward the target person (de Vignemont and Jacob, 2012, p. 307). We elaborate the benevolence we take empathy to imply below, sections "Social Understanding: Enactive and Reflective" and "Affective Intentionality: Bodily Resonance and Stoic Attachment."

state—the state of feeling empathy."). Empathy, on our account, not only involves understanding, but also an affective response to the state of the target person.

It should be clear by now that we cast the net wider than many other students of empathy. As the preceding remarks convey, we see no reason to restrict the concept to a purely epistemic other-comprehension or associating it narrowly with a particular type of other-comprehension. On our account, the phenomena described by Zahavi and de Vignemont are thus possible elements of empathy, but what sets empathy apart is the combination of its epistemic and affective elements. In what follows we shall try to develop an account that encompasses all these elements. Since the majority of work on empathy has focused on questions about the possibility of other-comprehension (Batson, 2009; Zaki and Ochsner, 2012), we shall start with a discussion of the epistemic aspects of empathy before we move on, in the section "Affective Intentionality: Bodily Resonance and Stoic Attachment" and "Discussion: Toward a Comprehensive Conception of Empathy," to discuss the affective responses involved in empathy and how these interrelate with its epistemic aspects.

Within discussions of social understanding "empathy" is often used more or less synonymously with the term "mindreading" as an answer to Batson's first question of how we can understand others[3]. This has traditionally been explained in terms of "theory of mind" (ToM), i.e., the basic insight that other people have minds and the abilities to reconstruct the inner life of others that follows upon this insight. It is controversial, however, whether these abilities are best understood as a type of inference, the so-called theory-theory, or as a type of simulation, the so-called simulation-theory (for good surveys of the debate, see Stueber, 2006, chap. 3; Gallagher, 2012b). Against this a growing number of researchers working within broadly phenomenological traditions have objected that these explanatory models unnecessarily complicate the matter by ignoring the extent to which the "inner" states of others are directly perceptible to us (Jensen and Moran, 2012; Gangopadhyay, 2014 provide good overviews of this critique).

This debate is sometimes described in terms of the different perspectives the individual approaches favor. Simulation-theory thus assigns the first-person perspective a fundamental significance in enabling social understanding. On this view, it is the ability to imagine yourself being in a certain state and then project this state onto others which explains the human ability to recognize what others are thinking and feeling (see e.g., Goldman, 2006; Stueber, 2006; de Vignemont and Jacob, 2012; Gallese and Sinigaglia, 2018). Theory-theory, on the other hand, likens the interpretation of others' behavior and inner states to the application of a scientific theory. On this view, social understanding relies on inferences about other people's behavior, facial expressions etc. based on a theoretical knowledge about how people normally think and feel when they behave in certain ways. This puts the empathizer in the role of an observer viewing the target person in a third-person perspective, drawing more or less objective inferences based on how s/he behaves (see e.g., Premack and Woodruff, 1978; Gopnik and Wellman, 1992; Carruthers and Smith, 1996). Finally, phenomenological and enactivist approaches to social understanding have stressed the fact that we normally engage with the thoughts and feelings of others in the context of direct social interaction where we experience the other in an immediate and interactive second-person perspective. Through the sensorimotor coordination and mutual attunement that characterizes direct, physical interaction the actions and intentions of others are immediately intelligible to us and this enables a direct or "primary" type of intersubjective understanding that does not rely on any inferences or oblique ascriptions regarding their mental states (see e.g., Gallagher, 2004, 2012b; Hutto, 2004; Zahavi, 2008; Fuchs and De Jaegher, 2009).

The basic assumption underlying the enactivist approach is therefore that "the mind of the other is not entirely hidden or private, but is given and manifest in the other person's embodied comportment," (Gallagher, 2004, 204; cf. Fuchs and De Jaegher, 2009, p. 469). As Gallagher points out it is only when "everyday second-person interactions break down, or when I have problems understanding the other person, I may engage in a specialized theoretical approach that appeals to third-person explanation or prediction. But such specialized cognitive approaches do not characterize our primary or everyday encounters with others" (Fuchs and De Jaegher, 2009, p. 202). Instead of relying only on traditional "cognitivist" approaches, social understanding therefore needs to be studied in terms of the enactive and embodied processes it involves (Fuchs and De Jaegher, 2009; Froese and Fuchs, 2012; Gallagher, 2012b; De Jaegher et al., 2017).

This enactive approach to social understanding has gained considerable traction within recent years and, we believe, it also offers the most promising starting point for an investigation of empathy. The most comprehensive, enactivist account of empathy to date is arguably the series of articles published individually and in collaboration with other leading researchers within the field by Thomas Fuchs, who draws extensively on both enactive and phenomenological discussions of social understanding (see most importantly Fuchs and De Jaegher, 2009; Froese and Fuchs, 2012; Fuchs and Koch, 2014; Fuchs, 2016, 2017).

Fuchs, like Zahavi, conceives of empathy as an immediate perception of the other's intentional state. He describes this as relying on "mutual incorporation" (Fuchs and De Jaegher, 2009; Fuchs, 2016) which constitutes a form of interaffectivity or "interbodily resonance," i.e., a mutual attunement or coordination in both movements and affections between interacting subjects (Fuchs and De Jaegher, 2009, p. 472–4; Fuchs and Koch, 2014, p. 5–7). Importantly, Fuchs here assumes that the immediate physical reaction to the other's expressions, e.g., the mirroring of a smile or the jolt away from an angry roar, is co-constitutive of the affective state we experience, in this case joy or fear. A feedback-cycle of affective responses can therefore arise: each participant picks up and reacts to

[3]Note that this question presupposes a more fundamental question about how we can know *that* others are thinking, sometimes referred to as "the problem of other minds." We shall not address this issue here.

the state of the other who then in turn reacts to her state (Fuchs and Koch, 2014).

The reciprocal and immediate nature of such interaction allows for participatory sense-making and an intimate attunement of emotional states. The interaction with others therefore provides an immediate grasp on their emotional states and intentions-in-action because we perceive and participate directly in them (Fuchs and De Jaegher, 2009; Fuchs, 2013; Fuchs and Koch, 2014). As Fuchs stresses, the sensorimotor coordination happens pre-reflectively and more or less instantaneously and the dynamics of this inter-affective system becomes "highly autonomous and are not directly controlled by the partners" (Froese and Fuchs, 2012, 213; see also Fuchs and De Jaegher, 2009; Fuchs, 2016). Fuchs therefore concludes: "Thus, emotions are not inner states that we experience only individually or that we have to decode in others, but primarily *shared states* that we experience through interbodily affection" (Fuchs and Koch, 2014, p. 7, their emphasis).

In contrast to de Vignemont's simulation-based account of empathy, Fuchs does not assume that the state of the empathizer and target person have to be the same. They are shared in the sense that they have their common origin in the dynamic interaction, but this can also involve complementary reactions. Also, we do not have to become aware of the shared state or explicitly ascribe it to the other in order to empathize. The understanding characteristic of this intercorporeality can of course be made explicit and reflective, but Fuchs emphasizes its fundamentally pre-reflective, implicit nature (Fuchs and De Jaegher, 2009; Froese and Fuchs, 2012; Fuchs, 2017). Supplemented by what Fuchs calls bodily memory, i.e., an implicit, pre-reflective memory which encapsulates both the interactive patterns of particular relationships and a more general knowledge of how to interact with others, this enables a fundamental kind of intersubjective understanding which Fuchs identifies as "primary empathy" (Fuchs, 2016, 2017).

Fuchs seems to assume that this basic form of empathy scales straight forwardly into more explicit and reflective attempts to understand others, like the ones subsumed under de Vignemont's definition of empathy, without the need for any further processes. He thus contrasts primary empathy with any attempt to reconstruct the inner states of others by means of logical inferences or simulation, which he calls "extended empathy" (Fuchs, 2013, 2017). According to Fuchs, these cognitive mechanisms are mainly relevant in cases of ambiguity or misunderstanding, whereas primary empathy sufficiently accounts for our everyday social interactions (Fuchs, 2017). Extended empathy appears to be a last resort or "deprived" version of primary empathy (Fuchs and De Jaegher, 2009, p. 472).

Fuchs' account of primary empathy thus ascribes both psychological and epistemic immediacy to the grasp of others' affective states provided through direct perception, i.e., it neither involves further psychological mechanisms (such as inference, projection or explicit ascription) nor does it require any further epistemic justification (as provided by e.g., ToM and "folk psychology"). His demotion of the cognitive mechanisms associated with extended empathy to a purely auxiliary or supplementary function, moreover, suggests that he takes this immediacy to characterize the majority of the social understanding involved in everyday interaction. This is a strong claim and, as we argue in the following section, it must be strongly modified.

While we take Fuchs' account of empathy with its basis in interbodily resonance and mutual incorporation to offer a promising way of capturing the embodied and enactive character of empathic engagements, the scope of the understanding provided by such engagements is in our view more limited than he appears to acknowledge. Based on a critical discussion of the forms and limits of social understanding informed by Schutz's phenomenological analysis of intersubjective meaning, the following section argues that even in everyday interaction empathic understanding involves a complex interplay between various processes, including both primary empathy and ToM. The fundamental significance of interbodily resonance and mutual incorporation for empathy, however, is not restricted to the epistemic aspects highlighted by Fuchs, but is equally, or perhaps even more closely, tied to their role in shaping our affective response to others.

SOCIAL UNDERSTANDING: ENACTIVE AND REFLECTIVE

Before we move on let us look more carefully at the account of embodied intentionality underlying Fuchs' account of primary empathy. On this account, empathy relies on our body's general capacity to resonate with its environment. This is brought out in his discussion of bodily resonance and embodied affectivity (Fuchs and Koch, 2014). As Fuchs stresses "bodily feelings should never be conceived as a mere by-product or add-on, distinct from the emotion as such, but as the *very medium* of affective intentionality." (Fuchs and Koch, 2014, p. 3 *sic*). By positing a distinctively bodily format for our affective intentionality Fuchs assigns the body a fundamental role in shaping how we perceive the world, allowing him to distinguish between a primary, embodied affectivity and the reflective processes that rely on it (cf. Goldman and de Vignemont, 2009 on this distinction between embodied and reflective processes).

The nature of bodily feelings, however, is twofold. Every feeling is simultaneously an affection and an expression. It registers an affective quality of the environment through e.g., visual, olfactory or tactile input, while the bodily changes this involves (facial, gestural, tensional etc.) also express our reaction to these "affective affordances" often including a tendency to move in certain ways, e.g., toward or away from an object (a tendency Fuchs dubs an "e-motion"). It is this registration-cum-reaction that Fuchs refers to by the broad term "emotion." Emotions, according to Fuchs, are thus the key to understanding intentional action: "Without emotions, the world would be without meaning or significance; nothing would attract or repel us and motivate us to act" (Fuchs and Koch, 2014, p. 2).

When we interact with other living subjects our intrabodily resonance becomes entwined in a reciprocal process of affection and expression: I pick up on your expressions and inevitably

react to them, if even in the slightest of ways, and my reaction in turn impacts you and elicits a new expression on your part. Like our intra-bodily resonance, this inter-bodily resonance is circular, but it extends beyond the individual to include all the interacting subjects thereby creating lesser or stronger feelings of "mutual incorporation" among them. This intercorporeality and sharing of states, as we saw above, are what enable empathic understanding on Fuchs account (Fuchs and Koch, 2014, p. 5–7).

As Fuchs and colleagues convincingly argue, the sensorimotor coordination and interbodily resonance involved in direct interaction thus provides an immediate access to the emotions and intentions of others in the sense of making these directly perceptible to us (for a review of recent studies into the neuronal bases of these mechanisms offering an interpretation in more simulationalist terms, see Gallese and Sinigaglia, 2018; cf. Stueber, 2006, p. 131–152). When we interact with others the majority of the movements, facial expressions, odors etc. they exhibit do not seem to require the inference of a hidden inner state in order for them to acquire meaning. In an important sense, your expression really is your emotion (recall the definition of bodily feelings as both affections and expressions). Thanks to the human ability to register and react to the bodily feelings of other living subjects their expressions therefore appear meaningful in the same immediate way that a piece of charcoal appears black and a stone hard.

This conception of interbodily resonance is capable of accounting for much of the interpersonal understanding involved in everyday encounters. It thus seems plausible that e.g., the yelling and tenseness of another person is directly perceived as "anger" rather than as indications allowing us to infer that the person in question is angry. But as we suggested above, there are also important limitations to such understanding. The basic contours of these limitations were brought out by Alfred Schutz in his penetrating analysis of social understanding (1967, esp. 20–38).

Schutz's point of departure is a critique of Scheler (1954) who, much like Fuchs, had pointed to the immediate, pre-reflective insight into the states of others offered by bodily expressions as a basis for empathy. Schutz, however, insists that this understanding is more limited than Scheler acknowledges (on this early phenomenological debate about empathy and Schutz' contribution to it, see Zahavi, 2010). First of all, not all expressions are also an attempt to express something, and even when this is the case, the agent's expressions only offer an indirect insight into their intentions: seeing someone who is yelling and raging, we immediately understand that the person is angry, but understanding their anger is more complicated. Grasping the full meaning of this state involves placing the agent's expressions within a broader context of meaning and interpreting them in that light. This is where things get murky (Schutz, 1967, p. 22–24).

To elucidate this interpretive process Schutz, building on Husserl's distinction between "that which is meant" (*Bedeutung*) and the act of "meaning" (*Bedeuten*), distinguishes between a level of "objective" meaning and one of "subjective" or "intended meaning" (*intendierter Sinn*) (Schutz, 1967, 33ff.). As he emphasizes the interpretive schemata we apply to the expressions of others (as well as our own) are to a large extent shared, intersubjective and in that sense "objective," but this intersubjectivity always presupposes an individual act of meaning-endowment (*Sinngebung*) (see further Schutz, 1967, chap. 2). You and I may therefore agree about the meaning of a given expression but we do so on the basis of distinct, and possibly differing, acts of meaning. My subjective experience will therefore necessarily differ from yours simply by virtue of it being mine and not yours.

A basic upshot of this distinction, as we interpret it, is that there are two levels on which you may strive to understand the meaning others experience: understanding it at the properly intersubjective level of identifying the meaning-content of their expressions, actions and intentions; and understanding how they experience this content, which remains essentially inaccessible, because this relies on an intentional operation of their consciousness that most often eludes even the awareness of the agent herself (Schutz, 1967, p. 34–37). The point of this distinction is not that social understanding is impossible, but simply that one should be aware that the meaning one is ascribing to others is of the intersubjective kind and as such a product of one's own intentionality—not the subjective meaning of the person one is trying to understand (Schutz, 1967, p. 30–31 and esp., 37–38).

We shall not go into any detail about the elaborate and sophisticated account of social understanding Schutz goes on to develop, but only offer a few selective observations (see Schutz, 1967, chap. 3; with Zahavi, 2010; León and Zahavi, 2016). For Schutz, the lived experience prior to any act of meaning-endowment is the privileged level of experience and we can crucially share this immediate experience of our existence with others through direct interaction, interlocking our intentionality with theirs and sensing their stream of consciousness. This immediate and pre-reflective other-comprehension (*Fremdverstehen*) is the closest one gets to the subjective meaning of others. The very act of meaning-endowment, by contrast, is reflective in the very basic sense that it imposes a certain meaning on our immediate perception. The moment we begin assigning meaning to the experience of others, we are therefore engaging in a private act of reflection rather than the special type of perception associated with true other-comprehension.

Corresponding to the sharing and interlocking of intentionality, which Schutz regarded as characteristic of true other-comprehension, Fuchs speaks of mutual incorporation as involving a decentering of our intentionality (Fuchs and De Jaegher, 2009, p. 476). Like Schutz, he assumes that this direct engagement offers an immediate and often quite intimate experience of the other's consciousness. Unlike Schutz, however, he seems to assume that it also grants a quite extensive, immediate understanding of the meaning they assign to their experience. This, Schutz would insist, is where we must distinguish between subjective and intersubjective meaning.

On a Schutzian analysis, what we share in cases of mutual incorporation is strictly speaking not our intentionality, in the sense of directly participating in each other's acts of meaning-endowment. Instead we attune and align our individual meaning-endowments and then share the meaning-content constituted

by these intentional operations. Thanks to the intersubjective character of the meaning-content each of us assign to our experiences, we are likely to assign similar content to many of the experiences we share during our interaction, thereby making your acts of meaning-endowment (including your emotions in Fuchs' broad sense) immediately intelligible to me. But this is due to a contingent overlap in the individual acts of meaning-endowment that each of us perform and the greater the difference in our past experience, cultural background etc., the greater the risk that our acts of meaning-endowment drift apart and I end up misunderstanding you.

Such misunderstandings are often quite subtle and of little practical importance to our social interaction and as such they pose no threat to Fuchs' claim that the immediate understanding derived from direct perception provides the predominant basis for everyday human interaction. The very negligibility of our misunderstandings, however, also points to a fundamental limitation in Fuchs' account of empathy: how can we know whether we really understand the other's experience or not? Since direct perception offers no access to the process of meaning-endowment, it provides no independent justification for our ascriptions of meaning-content to the experience of others and no means for detecting possible errors—except for the one's that incidentally disclose themselves in the course of our interaction. We are therefore liable to miss the often quite subtle differences between our own perspective and that of the other.

The "interbodily resonance" and "bodily memory" that Fuchs points to as enabling our direct perception of others' intentional states certainly provides some justification for the empathizer's assignment of meaning-content to the target's state (cf. Stueber, 2006, p. 142–147). Even though the human capacities to interact with and understand others in immediate and pre-reflective ways are indeed impressive, however, they can only take us so far. Consider a 3 year-old giving her father a picture she has painted. It is immensely plausible that due to their interactive history she possesses an "implicit," bodily knowledge of his reaction (Fuchs and De Jaegher, 2009; Fuchs, 2016; cf. Lyons-Ruth et al., 1998), that creates an immediate understanding and maybe even an anticipation of his reaction—her entire bodily posture reflects her enjoyment of the gratitude and compliments she is about to receive, even before the picture has left her hands. But all this—and this really is quite a bit of complex, social understanding—tells the child nothing about whether her father's reaction is caused by the picture, her act of generosity, or whether he is just being polite in order not to let her expectations down.

Fuchs never points to any mechanisms capable of providing such understanding. Empathy as he construes it may therefore provide us with a fairly reliable insight into what others experience, but arguably, his exclusive focus on recognitional capacities fails to acknowledge that we are not always just after the "what" of others' intentional states. Quite often the question driving our interest is how they came to see things or act this way. Moreover, these two types of understanding are plausibly seen as mutually informing: if my daughter gets the suspicion that I am dissimulating my joy, this conscious and reflective suspicion is likely to impact her immediate bodily reaction to my expressions of joy, regardless of the sincerity of my expressions (cf. Schmidsberger and Löffler-Stastka, 2018 who raise a similar point against Fuchs).

As opposed to the parallel assignment of the same content to a shared experience, we therefore suggest, the attempt to grasp the meaning-content of others, as this really appears to them, is a more ambitious task to the degree that it aims at understanding not just what the other intends, but also how and why they intend it. This involves a broadening in the scope of our other-directed intentionality from a sole focus on the meaning-content of the target's state to include the act of meaning-endowment producing that content.

Given that meaning-endowments are not directly observable, mostly eluding even the subjects own introspective "glance," reflecting on them is likely to include more than the direct perception of their effects. Such understanding, it would seem, relies on considerations about the context and processes informing the target's act of meaning-endowment as entailed by ToM and "folk psychology" (Stueber, 2006, chap. 4) or narrative competence (Gallagher, 2012a). We should therefore not, as suggested by Fuchs, conceive of these reflective, cognitive mechanisms as mere extensions or "deprived" versions of our immediate perceptual capacities, but as granting their own distinctive type of social understanding.

The very operation of trying to understand the meaning-endowments of others thus bears a close affinity to the act by which we focus attention on our own acts of meaning-endowment. Understanding e.g., why I reacted with fury in a given situation is not that different from understanding why you did so. Both operations necessarily focus on already elapsed (or future) experiences and both imply a shift from the engaged and immediate experience of the world to a sort of meta-perspective on how we/the other experience it. Reflection thus imposes a certain distance and loss of immediacy, but in contrast to the heuristic methods of direct perception, it also opens our understanding of the other to conscious evaluation and revision and thereby enables a dialogue between the interacting subjects.

In this sense, the reflective engagement with the experiences of others may be said to offer a deeper and more truly interpersonal understanding. Whereas the first-person experience of your intentionality is strictly speaking only accessible to you (with the possibilities of interpersonal resonance and shared meaning production opened up by direct interaction), it thus seems possible that others possess a deeper reflective understanding of your acts of meaning-endowment than you do yourself. After all, it is not uncommon that it takes an outside observer to make you aware that you are attributing a certain meaning to certain phenomena. In that instance the outside observer, in virtue of her correct hypothesis about your unconsciously experienced meaning, can reasonable be said to have possessed a deeper understanding of your intentional state than you did yourself.

This suggests that both with regard to self- and other-awareness we should recognize two distinct ways of understanding an intentional state: an immediate, perceptual grasp and a reflective interpretation [cf. Stueber's (2006) distinction between basic and reenactive empathy and Gallagher's (Gallagher, 2012a,b) distinction between "low-level" and "high-level" processes]. One might, we suggest, speak

of an "enactive" and a "reflective" understanding of others, thereby highlighting the immediate, interactive, and embedded character of the former, which derives from interbodily resonance, and the meta-perspective on the other's intentional states implied by the latter.

Empathy, as an other-awareness directed specifically at the affective states of others, aims at both these types of understanding, we take it, but to variable degrees depending on the context. We have pointed to Fuchs' analysis of interbodily resonance as providing a convincing account of the processes enabling enactive understanding, while we believe that reflective understanding to some degree relies on all the processes associated with ToM, folk psychology and narrative competence. The significance of these perceptual and interpretive processes in producing empathy, however, is not limited to the epistemic aspects we have focused on so far. They also have a significant impact on the affective attitude we adopt toward others.

Fuchs briefly touches on the other-directed, affective aspects of empathy in distinguishing empathic engagements from other, more objectifying types of interaction (Fuchs, 2013, p 660–661) but he never develops these ideas further. In the following section, we suggest that empathic engagements based on interbodily resonance are in fact characterized by a basic pro-attitude toward the other. We develop this suggestion with inspiration from the ancient Stoics and their theory of *oikeiōsis* or attachment. Based on a theory of perception and intentional action that is closely similar to Fuchs' concepts of bodily resonance and affective intentionality, these ancient Greek philosophers developed an analysis of interpersonal relations and prosocial motivation that, so we suggest, can usefully supplement Fuchs' concept of primary empathy and help us understand the affective aspects of this emotional phenomenon.

AFFECTIVE INTENTIONALITY: BODILY RESONANCE AND STOIC ATTACHMENT

The Stoic school established itself as an influential intellectual tradition under the leadership of Zeno, Cleanthes, and Chrysippus in the generations just after Plato and Aristotle and it came to dominate many of the philosophical debates throughout the Hellenistic world well into the second century CE. As part of their overarching project of figuring out how human beings can come to live in perfect harmony with themselves and the world around them, these highly original thinkers developed a sophisticated analysis of how humans orient themselves in the world, their so-called theory of *oikeiōsis*. The structuring idea of this theory is an ability to recognize different things as belonging or proper to oneself (gr. *oikeion*). This, according to the Stoics, produces a sense of attachment (*oikeiōsis*) that structures our perception of the world in much the same way as we perceive our environment in terms of "affective affordances," i.e., the affective qualities it possesses for us, on Fuchs' account of affective intentionality.

Based on rather crude, but for their time groundbreaking physical and medical theories, the Stoics describe the perceptual processes underlying this intentionality in highly visceral terms as a consequence of our material soul rubbing against the body, with which it is coextensive and thoroughly blended:

> For the soul extends outward with an expansion and strikes all the parts of the body, since it is also mixed with all of them, and when it strikes them it is struck back in turn. For the body too offers resistance, just like the soul: and the affection ends up being simultaneously characterized by pressure and counterpressure [Hierocles, 2009, p. 13 (IV.44–49)].

The affections (*pathē*, a term that also means feelings or emotions) produced through this interaction of body and soul are also described as a reciprocal co-affection between the two—*sympatheia* in ancient Greek [Hierocles, 2009, p .11 (IV.1–22)]. It produces not only an awareness of the various parts of our body but also an instinctual grasp (*antilēpsis*) of their functions and needs [Hierocles, 2009, p. 3–9 (I.1-III.19); and Seneca, 2007, p. 85–89 (121.5–6 and 22–24); with Long, 1996; Brittain, 2002, p. 266–269; Klein, 2016, 172–176]. This, in turn, explains how we perceive external objects. When we encounter something, it affects our body and the soul registers this affection making us aware both of the object and how it affects us [Hierocles, 2009, p. 16 (VI.1–9)]. Depending on whether this affection is perceived to fit with our immediate constitution and needs, it will appear either proper (*oikeion*), foreign (*allotrion*), or neither to us [Hierocles, 2009, p. 9–11 (III.20–54); and Seneca, 2007, 88–89 (121.17–21); with Brittain, 2002, p. 269–271].

The continuous awareness or co-perception (*synaisthesis*) of ourselves also results in a feeling of attachment (*oikeiōsis*) and a basic disposition to show concern (*tērein*) for ourselves—we recognize the various aspects of our constitution as proper to us and therefore start caring for them—and this in turn, is said to produce impulses to pursue things that are *oikeion* and avoid their opposites [Laertius, 1925, p. 193 (VII.85); Cicero, 2001, p. 69–70 (III.16–18); Seneca, 2007, p. 85–89; Hierocles, 2009, 17–21 (VI.28–VII.50); with Klein, 2016, p. 165–178]. An impulse (*hormē*), on the Stoic account, is thus the psychic event causing an animal's body to move (Inwood, 1985, p. 42–101; Brennan, 2003; Graver, 2007, p. 24–28). For instance, when I am hungry and see a banana, I recognize the banana as suitable nourishment, i.e., as *oikeion* to me, and eating the banana therefore strikes me as appropriate (*kathēkon*) resulting in an impulse to eat the banana.

On the Stoic view, there are thus four psychological mechanisms underlying every action:

1. the perception of something as *oikeion*, *allotrion* or neither through bodily affections/feelings (corresponding to the embodied appraisal of affective affordances on Fuchs' account of affective intentionality)
2. the natural tendency to feel attachment toward things we recognize as *oikeion*
3. the disposition to show concern (*tērein*) for ourselves and, derivatively, the things we feel attached to
4. the impulse to move (corresponding to Fuchs' e-motions).[4]

[4]A note of clarification: The analogy that we suggest is between Fuchs' bodily feelings and the so-called "hormetic," i.e., impulse-producing, impressions (*hormētikai phantasiai*), not the subspecies of misguided reactions that the Stoics provocatively called emotions (Graver, 2007, p. 15–34). The Stoics thus analyzed the reaction to multisensory stimuli into two distinct psychic events: impressions,

Despite the dualistic coloring of their terminology[5] the Stoic concept of a *sympatheia* between body and soul therefore bears a close resemblance to Fuchs' concept of bodily resonance: Both make bodily feelings the medium of affective intentionality and both see these affections as entailing an awareness both of the object affecting you and of yourself being affected, accompanied by a tendency to act. The Stoics, however, add the crucial elements of self-attachment and -concern.

Fuchs' speaks more neutrally of the self-reference involved in affective intentionality (Fuchs and Koch, 2014; cf. Slaby and Stephan, 2008). This would seem to involve some kind of basic pro-attitude toward oneself—a recognition of myself as "me" and a wish to preserve this self, similar to the self-perception, -attachment, and -concern posited by the Stoics—but this remains somewhat unclear: "They (sc. Emotions) always imply a particular relation to the feeling subject in its very core: through emotions, I experience *how it is for me* to be in this or that situation. To be afraid of an approaching lion (world-reference) means at the same time being afraid for oneself (self-reference). Each emotion, thus, implies the two poles of feeling *something* and feeling *oneself* as inextricably bound together." (Fuchs and Koch, 2014, p. 3, their emphasis). "Feeling oneself" here seems to include both a perception of oneself ("I experience *how it is for me*") and a concern for oneself ("being afraid for oneself"). This ambiguity between perception/understanding and concern pervades Fuchs' account of bodily resonance and to some extent it is inherent in the very words "feeling" and "affection." The Stoics, on the other hand, effectively dissolve this ambiguity by analyzing the co-perception of the self into the three closely linked, but distinct elements of perception, attachment, and concern: Through my bodily feelings I perceive myself being affected, I recognize this self as something I feel attached to, and therefore I care about what happens to me. The Stoic theory of attachment can therefore helpfully supplement Fuchs' phenomenological analysis of "incorporation" by providing a terminology for the relational aspects of this way of "having a world" (Merleau-Ponty, 1962, p. 169) and the affective value it supposes us to ascribe to our own body and its surroundings.

As Fuchs stresses incorporation is "a pervasive characteristic of the lived body, which always transcends itself and partly merges with the environment." (Fuchs and De Jaegher, 2009, p. 472). This includes the unidirectional integration of an object into the sensorimotor schema of the lived body which allows us to interact and coordinate with it as e.g., the blind person incorporates a cane; but it also covers the "intercorporeality" or "mutual incorporation" made possible through the reciprocal interbodily resonance between two living subjects (Fuchs and De Jaegher, 2009, cf. Merleau-Ponty, 1962). The latter type of incorporation, as Fuchs points out, enables us to perceive and respond immediately to the affective states of others, but whereas he goes into considerable detail about the kind of perception he takes to be characteristic of such direct empathy, his account of the affective responses appropriate to empathy remains underdeveloped. The richer Stoic concept of attachment, however, offers the conceptual means for spelling out and clearly delineating the scope of these aspects of empathy.

Not unlike Fuchs, the Stoics held that the *sympatheia* characterizing the interaction of body and soul extends beyond your own organism. In principle, it extends to the entire universe which for them is a compound of body (or matter, *hyle*) and soul (or fiery air, *pneuma*) just like the human organism (Brouwer, 2015). On this basis they seem to have posited some limited possibilities of sharing affective states with other living beings and the world at large (see Brouwer, 2015 for a more detailed discussion), but in contrast to Fuchs, they never seem to have explored the possible epistemic implications of this doctrine[6]. Instead, their analysis of social interaction, like that of self-perception, is carried out in terms of attachment and concern.

While our most fundamental feelings of attachment pertain to ourselves and our constitution, this attitude extends well beyond the borders of our own body to include anything that we recognize as *oikeion* to us. Occasionally this attitude was analyzed into the basic well-disposed (*eunoētikē*) attachment toward ourselves which determines the selective (*hairetikē*) attachment we feel toward external things and the affectionate (*sterktikē*) attachment we feel toward other people (Hierocles, 2009, p. 25 [IX.1-10]; cf. Anonymus, 1995, VII–VIII).

The fact that the Stoics used the same term to describe such different relations suggests that they saw a basic similarity in the character of those relations. The surviving Stoic texts, however, do not offer any univocal explanation of what it is that makes something qualify as *oikeion*—apart from the formal answer that everything that is part of or somehow agrees with our constitution is *oikeion* to us (cf. Klein, 2016)—but the etymology of the term provides some fascinating hints about how the Stoics envisioned the relations it describes. *Oikeion* thus derives from *oikos* ("house" or "household") and most directly it describes the relation one has toward the people and things one was brought up with. By extension, it came to apply to aspects of one's own

defined as an "alteration" or "affection (*pathos*) that happens in the soul and shows in itself the thing that produced it" [Long and Sedley, 1987, p. 237 (39B, Aëtius 4.12.2); with Brittain, 2002, p. 259–261]; and impulses (*hormē*), defined as a "movement" in the soul, which follows upon these affections when they are productive of impulses (Inwood, 1985, p. 42–66; Brennan, 2003). When this reaction is appropriate, it results in a "proper act" (*kathēkon*), when it is inappropriate, in an emotion (*pathos* again). In adults, these reactions presuppose a reflective act of "assent" to the hormetic impression (whereas in children and animals it follows directly on the impressions), but besides these (at least potentially) reflective reactions the Stoics also recognized a set of pre-reflective and therefore uncontrollable reactions called pre-emotions (*propatheiai*), i.e., bodily reactions which are independent of our assent (Graver, 2007, p. 85–108).

Parallel to Fuchs, we suggest, the Stoics thus see bodily affections/feelings as constituting an embodied appraisal of the environment in terms of *oikeion, allotrion* or neutral, which entails both an affection (the impression) and expression (the pre-emotion) caused by the external stimuli, while the impulses following upon this appraisal correspond to Fuchs' concept of a movement tendency or "e-motion."

[5]In overt rejection of Platonic dualism, the Stoics in fact stress the holistic unity of body, soul and intellectual mind (Gill, 2006). Cf. Fuchs' rejection of "the Cartesian framework" of traditional cognitive science (e.g., Fuchs and De Jaegher, 2009, p. 468; Fuchs, 2017, p. 27–30).

[6]The Stoics were e.g., (in)famous for claiming that in the ideal case one would be able to feel the effect of another person stretching out her finger at the other end of the world and share the joy of this movement if it was carried out well [Plutarch, 1976, p. 730–731 (Comm. not. 1068F)]. This, however, would be a case of sharing (in Scheler's sense of feeling the same due to the same event) and therefore involves no social understanding.

organism. Its Stoic meaning is therefore perhaps best captured by the English "proper" or "akin" (see further Pembroke, 1971). By contrast, the word *idios* ("personal, private, peculiar to you") more directly describes something that belongs to you as opposed to others[7]. The choice of the word *oikeion* to describe the way we perceive our own organism may thus be taken to reflect a view of the subject as deeply embedded in a social context and community rather than a more or less isolated individual[8].

Similar to the lived body in phenomenological theories of perception (see e.g., Merleau-Ponty, 1962; Gallagher, 2005, chap. 1), the "self" (in the Stoic sense of the mind-cum-body also referred to as our "constitution") is the pivotal point of our intentionality. The attachment and concern we feel toward this self furthermore explain the affective value we attribute to the things we perceive and our motivations to care for, pursue or avoid them. As for Fuchs, this perceptual process is pre-reflective and non-verbal. Our perceptions can be verbalized and reflected upon, but they happen through the medium of bodily feelings or affections[9].

On this view, the attribution of affective value to something, and thereby your concern for it, depends on the establishment of a relation between that object or person and your "self"—on you seeing them as part of or at least relevant to who you are and what you want. But this self is not a stable or monolithic thing. It is subject to constant change and re-configuration during a lifetime, and with it the things and people we recognize as *oikeion* change. The image of the household, at the root of the concept of attachment (or appropriation, as *oikeiōsis* is also frequently translated), is suggestive. Just like the members of your household, your furniture etc. change, and with them the constitution of your household, so your "self" and the things you recognize as belonging to it changes. The late Stoic Hierocles, (2009, p. 91–93 [apud Stob. 4.84.23]; cf. Cicero, 1991, 21–25 [I.50–60]) thus portrays the social world of every human being as structured in circles of ever more attenuated bonds of attachments, from your own body over your closest kin to the most remote stranger.

On its most convincing interpretation, the Stoic theory of attachment therefore reflects what Algra (2003) has called a model of "extended partiality." On this view, the attachment toward one's own organism may—at least at the outset of human development—represent a privileged relation, but in principle it is no different from the attachment we feel toward other things that are within our sphere of interest and concern. Attachment, on the Stoic conception, thus transcends the boundaries of the individual agent bringing a broader range of things within our sphere of concern and making them the co-reference point of our actions. It constitutes an extension of our affective intentionality, not only making the other affectively relevant to us but effectively including her interests in our sphere of concern. This model of extended partiality fascinatingly cuts across the entrenched modern distinction between egoistic and altruistic motivations and perforates the borders between "self" and "other" in ways that are suggestive of many modern, enactive, and phenomenological accounts of intersubjectivity and "intercorporeality"—including Fuchs'. The Stoic model of extended partiality, we suggest, therefore provides a promising way of describing the shift in motivations that accompanies the decentering of our intentionality posited by his account.

To sum up, we have argued that in order to fully grasp the significance of interbodily resonance and mutual incorporation in enabling empathic, social engagements one must take into account both its epistemic and affective aspects. We have suggested that the latter be analyzed along the lines of the Stoic theory of attachment and concern. On this account, empathy can be seen to imply a basic pro-attitude toward the other, a perception of the other that presents her as a proper object for our attachment. This, in turn, produces a concern that can generate the prosocial motivation to help her in particular circumstances or a more general attitude of extended partiality toward her. In the following section, we shall relate this model of empathic motivation to the modern debate about the relation between empathy and altruism.

DISCUSSION: TOWARD A COMPREHENSIVE CONCEPTION OF EMPATHY

Having examined both the epistemic and affective aspects of interbodily resonance in some detail, delineating its plausible effects in enabling social understanding and prosocial attitudes, we may now proceed to develop our attempt at a comprehensive account of empathy.

In the section "Empathy: Cognitive, Emotional, Primary, and Extended," we examined the embodied basis of social understanding in interbodily resonance but argued for the need to acknowledge both a reflective and an enactive type of social understanding. As maintained by Fuchs, enactive understanding quite plausibly supplies the necessary prerequisites for developing ToM and these two ways of engaging with the intentions and perspectives of others remain tightly integrated and mutually informing. Nonetheless, they are different capacities, accounting for different kinds of understanding and empathic engagement: enactive understanding offers an immediate access to the meaning-content of intentional states, while reflective understanding reasons about the subjective meaning and acts of meaning-endowments underlying such states.

Following Fuchs, we see interbodily resonance as fundamental both to direct empathic engagement and to the acquisition of empathic capacities, but with inspiration from the ancient

[7] *Idios* is the etymological root of the, often derogatory, term *idiōtēs*, which signifies the private citizen as opposed to the citizen actively taking part in the duties of the political community. Interestingly, we find [Plato (1997), 1,089 (462b)] coining the word *idiōsis* ("privatization" from *idios*, parallel to *oikeiōsis* from *oikeios*), contrasting it with *koinōnia,* "community."

[8] The Stoics were of course picking up on already established general and philosophical usages of the term *oikeion* (although they did to all appearances use the term far more systematically than any of their predecessors and thereby probably helped extend its meanings). To some degree the communal mindset we ascribe to them here was thus part of their cultural background. For a good discussion of how this mind set is reflected in Stoic social philosophy, see Algra (2003).

[9] Whereas Fuchs conceives of bodily feelings as non-representational (see e.g., Froese and Fuchs, 2012, p. 207, 213), however, this does not seem to be the case with the Stoics (Sorabji, 1993, p. 20–28; Brittain, 2002).

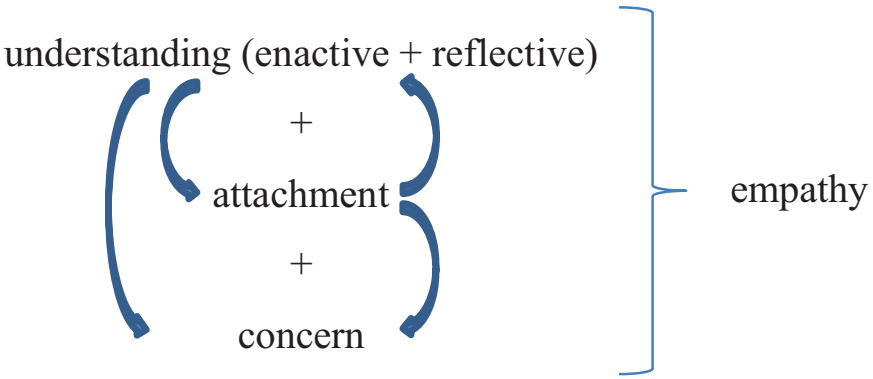

FIGURE 1 | The distinctive elements of empathy. This figure shows empathy as involving social understanding of both the enactive and reflective type, a sense of attachment toward to other and a concern for them that can motivate the empathizer to undertake action in the interest of the target person. Other-concern, we take it, arises on the basis of understanding, attachment, or both (as indicated by the downward arrows in the figure). The arrows between the two latter states indicate that these are most plausibly seen as mutually informing and re-enforcing (attachment can produce an increased propensity to engage in attempts to understand the other, just like an understanding of the other can produce attachment toward them).

Stoics we have argued that this type of interaction also creates feelings of attachment and concern. On our account, this pro-attitude toward others characterizes any empathic engagement and distinguishes such engagements from more objectifying, disengaged, or "cold-blooded" instances of social understanding. As we argued in the section "Social Understanding: Enactive and Reflective" above, attachment and the entailed recognition of the other as a proper object of concern furthermore suggests a conceptualization of the prosocial behavior often associated with empathy in terms of extended partiality.

Based on these discussions we believe that the relation between the epistemic, affective, and motivational aspects of empathic engagements as in **Figure 1**.

On this model, empathy is a complex emotional phenomenon that involves three distinct other-related states allowing for varying degrees of intensity and shifting interrelations: an understanding of the other's affective state, a feeling of attachment toward the other, and a concern for them. It is important to note that the individual processes underlying these states can of course also operate in isolation or appear in other constellations: one can understand the state that someone else is in without feeling neither attachment nor concern toward them (cf. Zahavi's expert torturer), just like understanding alone can produce concern without any elements of attachment (cf. Bloom's 2017a, b, concept of "rational compassion"). Similarly, one can feel attachment toward someone else without even trying to understand them, and this can likewise produce concern. What we take to be distinctive of empathy, however, is the co-presence and interplay of these processes and states.

Rather than narrowing the scope of empathy to the imaginative simulation of the other's affective state, as suggested by de Vignemont and colleagues, or to the immediately perceptual and purely epistemic engagement with others' states, as suggested by Zahavi, we thus define empathy in virtue of the co-presence of its constitutive elements. Empathy, on our account, therefore differs from mindreading more broadly by implying an attachment and concern for the other. On this account, the mindreading of Zahavi's expert torturer does not count as empathy; just like her good twin, the rationally compassionate helper who acts out of a fully detached appreciation of someone else's needs and an impartial, prosocial motivation, does not empathize. Also, we take empathy to be specifically concerned with the affective states of others (or emotions in Fuchs' broad sense) and only incidentally with non-valenced beliefs and perceptual states.

On the other hand, empathy differs from purely affective types of experiential sharing, like emotional contagion or a collective reaction to some event, by entailing an understanding of the other's affective state[10]. This epistemic element also distinguishes empathy from more general feelings of pity or sympathy, but if sympathy is conceived broadly as just any benevolent emotional response to others' affective states, empathy can of course be considered a type of sympathy[11]. Likewise, empathy differs from the mere feeling of concern, that is an element of it, and the motivation to help, that is a frequent consequence of it. On this account, empathy can therefore both be seen as a distinctively affective type of mindreading and a distinctively epistemic type of sympathy.

The distinctive feature of our account is thus the integration of understanding, attachment, and concern as constitutive parts of empathy. This conception of empathy is comprehensive in the sense that it encompasses the epistemic, affective and motivational aspects of empathy. It thereby addresses both of Batson's questions regarding empathy and insists that a proper answer to either requires us to consider the other as well. This is done on the grounds that social understanding, attachment

[10]Note, however, that emotional contagion and sharing, on our account, can lead to empathy *if* they bring about an awareness that the other is experiencing a similar affective state *and* this awareness is accompanied by a feeling of attachment and concern toward them.

[11]Sympathy carries shifting meanings, many of which overlap with the concept of empathy, but often it simply appears as an umbrella term for benevolent emotional responses to others (for a good overview, see Stueber, 2006, p. 28–31).

and concern appear to correlate in ways that elude us when studied in isolation.

Our working definition tries to convey this complexity, in terms that stay neutral with regard to the exact nature of the processes and mechanisms involved, as *a benevolent engagement with the affective states of others which provides us with a grasp of their state and produces an affective response within our bodies*. As mentioned in the section "Empathy: Cognitive, Emotional, Primary, and Extended" above, recent philosophical and empirical debates on empathy have generally focused on questions regarding the scope and importance of different psychological processes in enabling social understanding. Proponents of rivaling accounts, however, tend to agree that a full account of social understanding is likely to involve some combination of all the processes under consideration (Stueber, 2006; O'Shea, 2012; Fuchs, 2017; Gallese and Sinigaglia, 2018; Fernandez and Zahavi, 2020). By merely stipulating that empathy "provides a grasp of the other's state" our definition allows for different construals of how these processes figure and interact in empathy. In order to clarify the scope of the different proposals, however, we have suggested to distinguish clearly between the attainment of enactive and reflective understanding.

With regard to the former type of understanding, we have pointed to Fuchs' concept of interbodily resonance and affective intentionality as a convincing way to capture the embodied and enactive character of these processes; but there are also serious attempts to explain this in simulationalist terms (Gallese, 2003; Gallese and Sinigaglia, 2018) just like theory-theory approaches may incorporate many of the objections leveled against it by proponents of phenomenological and 4E approaches (O'Shea, 2012). Our hope is that the account offered here may help advance this debate as it relates to empathy by providing a pluralistic, conceptual framework for studying the epistemic and affective aspects of the relevant processes in combination thereby allowing for a more holistic assessment of their significance in enabling empathic engagements.

We thus take the processes producing empathic understanding and attachment to be mutually informing. As Cialdini et al. (1997) have pointed out an empathic concern deriving from perspective taking, on the one hand, and feelings of oneness, on the other, are likely to arise in the same contexts and they appear sensitive to many of the same factors (relationship closeness, severity of need etc.). This indicates that empathic understanding and attachment are closely correlated and we have pointed to the processes involved in interbodily resonance as important factors in producing both. It is likewise plausible to assume that empathic understanding can create or deepen feelings of attachment, just like feelings of attachment are likely to make you more prone to engage in an attempt to understand the other person (cf. Batson and Shaw, 1991; Cialdini et al., 1997).

Attachment has a long history as a term for describing the intimate bonds between a child and its closest caregivers (see e.g., the seminal paper by Bowlby, 1958) but it has not been used systematically to describe wider social bonds (cf. Batson and Shaw, 1991, p. 113). We have pointed to the Stoic theory of attachment as providing a concise conceptual framework for such an analysis of our social relations. On this theory, our social bonds rely on a proneness to see other people as *oikeion* to us and feel attachment and concern toward them, parallel to the way we perceive our own body and certain aspects of our environment as "belonging" to us. In extension of Fuchs' adoption of Gibson's (1979) theory of affordances, one might speak of "empathic affordances" and work toward an understanding of the factors informing this social aspect of our affective intentionality (factors such as kinship, similarity in cultural background, racial prejudice etc. but, perhaps even more importantly, purely bodily factors such as skin-to-skin-contact (Ciaunica, 2017), olfactory intake etc.).

Both the phenomenological concept of incorporation and the Stoic concept of attachment furthermore suggest that our perception of others is in some respects parallel to the way we perceive and relate to our own body. These proprioceptive processes have received increasing attention within contemporary psychology and neuroscience under the umbrella term "Body Ownership" (BO) covering studies of how we attach to/detach from our own limbs and bodies. The quantitative and qualitative studies of BO have highlighted how the sensation of something being my "own" body is dependent on visual and tactile intake and is therefore highly sensitive to manipulation of this intake—even to the degree that such manipulation can create the temporary, illusory adoption of a foreign limb or even an entire body as one's own (Botvinick and Cohen, 1998; Petkova et al., 2011; Maselli and Slater, 2013; Guterstam et al., 2015).

An important future avenue for empirical studies of empathy, suggested by our account and to some extent already being pursued, is to transfer the insights gained in BO-scholarship about the role of multisensory stimulation in producing feelings of ownership/attachment toward our own body to the intersubjective sphere, on the assumption that they play similar, causal roles in determining how we relate to the bodies of others (cf. Bertrand et al., 2018). Indeed, there is already tentative evidence that adopting another limb (or avatar body) that has a different skin tone also reduces implicit racial biases (Farmer et al., 2012, 2014; Maister et al., 2013; Peck et al., 2013; Farmer and Maister, 2017). The authors of these studies argue that looking through another body and feeling body ownership to it creates a kind of kinship which in turn lowers out-grouping tendencies. These studies provide examples of how even an illusory and temporary perception of someone else as *oikeion*—literally perceiving a foreign limb or body as belonging to you—impacts our attitude toward them, thereby raising important questions about the role of (perceived) similarity or relatedness in creating feelings of attachment, concern, and ultimately empathy, and the importance of pre-reflective, bodily factors in shaping such (a)kinship-perception.

Another important aspect of the Stoic account is the model it provides for describing the prosocial motivation associated with empathy. As brought out by the considerations above, the prosocial motivation that derives from feelings of attachment is inherently partial. It rests on the inclusion of others into your in-group and only thereby into your sphere of concern: the two of us,

we Europeans, we humans etc. By contrast, some theorists have assumed that empathy can inspire genuinely altruistic action, i.e., an ultimately selfless motivation to promote the wellbeing of the other induced through perspective taking (Batson and Shaw, 1991). As Cialdini et al. (1997) have convincingly shown, however, feelings of oneness appear to have a far more decisive impact on our decisions to help, suggesting that the potential effects of empathic understanding on prosocial motivation are mediated by such feelings.

This is conveyed by our insistence that empathy necessarily involves a feeling of attachment toward the other and we have described the resulting type of motivation as an extended partiality, i.e., the inclusion of the other into our sphere of concern based on a perception of the other as *oikeion* to us. Herein lies the strengths and limitations of empathy as a motivational force: it brings the affective states and needs of its target into sharp focus, thereby inevitably forcing those of others into the background (cf. Bloom, 2017a, b). Rather than mobilizing us "against empathy," however, this remarkable capacity of human beings to extend their partiality in our view leaves plenty of room for optimism about the moral value of empathy.

CONCLUSION

After a brief review of recent debates about empathy this paper set out to examine and expand upon the enactivist account offered by Thomas Fuchs. Through a critical discussion of his conception of bodily resonance as providing the basis for an immediate understanding of other's intentional states, we have argued for the need to acknowledge the equal importance of reflective, social understanding in everyday, human interaction and allow for different processes to provide the predominant foundations of these types of understanding.

As a distinctively affective, other-directed intentionality, however, empathy does not merely consist in understanding the affective states of others. It also involves an affective response toward the other and their situation (as the empathizer perceives this). This response, we have suggested, can be analyzed along the lines of the Stoic theory of *oikeiōsis* as a feeling of attachment and concern toward the other that arises on the basis of interbodily resonance along with other, more reflective processes that makes the other appear *oikeion* or (a)kin to us. This inclusion of the other in one's sphere of concern produces, in turn, a prosocial motivation that can best be described as an extended partiality.

Based on these philosophical reflections we have proposed an account of empathy as a complex emotional phenomenon comprising epistemic, affective, and motivational elements and we have briefly related the Stoic concepts we introduced in order to describe its affective and motivational elements to existing concepts within current debates about Body Ownership and altruistic motivation. Our approach has been shamelessly eclectic—integrating concepts from ancient Stoicism, modern phenomenology, psychology, and neuroscience—but the result, we submit, is an account of empathy that acknowledges the complexities of this other-directed and inherently partial way of engaging with others and can thereby help increase conceptual clarity across the interdisciplinary field of empathy studies.

DATA AVAILABILITY STATEMENT

All datasets generated for this study are included in the article/supplementary material.

AUTHOR CONTRIBUTIONS

The ideas for this article along with a first draft were developed in close collaboration between the authors. The final version was written by TS with comments and suggestions from KB. Both authors contributed to the article and approved the submitted version.

FUNDING

Work on this manuscript was made possible through the generous support of the Deutsche Forschungsgemeinschaft (DFG) awarded to TS through the research training group "Philosophy, Science and the Sciences" at the Humboldt University, Berlin, who also covered the publication costs.

ACKNOWLEDGMENTS

We are very grateful toward the reviewers for their acute comments and immensely useful suggestions for improvements. We would also like to thank the audience at the philosophical colloquium at Humboldt University, Berlin, and Maxie Schulte for their comments on earlier versions of this manuscript.

REFERENCES

Algra, K. (2003). The mechanism of social appropriation and its role in hellenistic ethics. *Oxf. Stud. Anc. Philos.* 25, 265–296.

Anonymus (1995). "Commentarium in platonis theaetetum," in *Corpus Dei Papiri Filosofici Greci e Latini*, Vol. 3, eds D. Sedley and G. Bastianini (Leuven: Leuven University Press) 221–562.

Batson, C. D. (2009). "These things called empathy: eight related but distinct phenomena," in *The Social Neuroscience of Empathy*, eds J. Decety and W. Ickes (Cambridge, MA : The MIT Press).

Batson, C. D., and Shaw, L. L. (1991). Evidence for altruism: toward a pluralism of prosocial motives. *Psychol. Inq.* 2, 107–122. doi: 10.1207/s15327965pli0202_1

Bertrand, P., Guegan, J., Robieux, L., McCall, C. A., and Zenasni, F. (2018). Learning empathy through virtual reality: multiple strategies for training empathy-related abilities using body ownership illusions in embodied virtual reality. *Front. Robot. AI* 5:26. doi: 10.3389/frobt.2018.00026

Bloom, P. (2017a). *Against Empathy: The Case for Rational Compassion*. New York, NY: Random House.

Bloom, P. (2017b). Empathy and its discontents. *Trends Cogn. Sci.* 21, 24–31. doi: 10.1016/j.tics.2016.11.004

Botvinick, M., and Cohen, J. (1998). Rubber hands 'Feel' touch that eyes see. *Nature* 391, 756–756. doi: 10.1038/35784

Bowlby, J. (1958). The nature of the child's tie to his mother. *Int. J. Psycho Anal.* 39, 350–373.

Brennan, T. (2003). "Stoic moral psychology," in *The Cambridge Companion to the Stoics*, ed. B. Inwood (Cambridge: Cambridge University Press).

Brittain, C. (2002). Non-rational perception in the stoics and augustine. *Oxf. Stud. Anc. Philos.* 22, 253–308.

Brouwer, R. (2015). "Stoic sympathy," in *Sympathy: A History*, ed. E. Schliesser (Oxford: Oxford University Press), 15–35. doi: 10.1093/acprof:oso/9780199928873.003.0002

Carruthers, P., and Smith, P. K. (1996). *Theories of Theories of Mind*. Cambridge: Cambridge University Press.

Cialdini, R. B., Brown, S. L., Lewis, B. P., Luce, C., and Neuberg, S. L. (1997). Reinterpreting the empathy–altruism relationship: when one into one equals oneness. *J. Pers. Soc. Psychol.* 73:481. doi: 10.1037/0022-3514.73.3.481

Ciaunica, A. (2017). The 'Meeting of Bodies' – empathy and basic forms of shared experiences. *Topoi* 38, 185–195. doi: 10.1007/s11245-017-9500-x

Cicero, M. T. (1991). in *On Duties*, ed. M. T. Griffin trans. E. M. Atkins (Cambridge: Cambridge University Press). doi: 10.1007/s11245-017-9500-x

Cicero, M. T. (2001). in *On Moral Ends*, ed. J. Annas trans. R. Woolf (Cambridge: Cambridge University Press). doi: 10.1007/s11245-017-9500-x

De Jaegher, H., Pieper, B., Clénin, D., and Fuchs, T. (2017). Grasping intersubjectivity: an invitation to embody social interaction research. *Phenomenol. Cogn. Sci.* 16, 491–523. doi: 10.1007/s11097-016-9469-8

de Vignemont, F., and Jacob, P. (2012). What is it like to feel another's pain? *Philos. Sci.* 79, 295–316.

de Vignemont, F., and Singer, T. (2006). The empathic brain: how, when and why? *Trends Cogn. Sci.* 10, 435–441. doi: 10.1016/j.tics.2006.08.008

Farmer, H., and Maister, L. (2017). Putting ourselves in another's skin: using the plasticity of Self-perception to enhance empathy and decrease prejudice. *Soc. Justice Res.* 30, 323–354. doi: 10.1007/s11211-017-0294-1

Farmer, H., Maister, L., and Tsakiris, M. (2014). Change my body, change my mind: the effects of illusory ownership of an outgroup hand on implicit attitudes toward that outgroup. *Front. Psychol.* 4:1016. doi: 10.3389/fpsyg.2013.01016

Farmer, H., Tajadura-Jiménez, A., and Tsakiris, M. (2012). Beyond the colour of my skin: how skin colour affects the sense of body-ownership. *Conscious. Cogn.* 21, 1242–1256. doi: 10.1016/j.concog.2012.04.011

Fernandez, A. V., and Zahavi, D. (2020). Basic empathy: developing the concept of empathy from the ground up. *Int. J. Nurs. Stud.* 110:103695. doi: 10.1016/j.ijnurstu.2020.103695

Froese, T., and Fuchs, T. (2012). The extended body: a case study in the neurophenomenology of social interaction. *Phenomenol. Cogn. Sci.* 11, 205–235. doi: 10.1007/s11097-012-9254-2

Fuchs, T. (2013). The phenomenology and development of social perspectives. *Phenomenol. Cogn. Sci.* 12, 655–683. doi: 10.1007/s11097-012-9267-x

Fuchs, T. (2016). "Intercorporeality and interaffectivity," in *Intercorporeality: Emerging Socialities in Interaction*, eds C. Meyer, J. Streeck, and J. S. Jordan (Oxford: Oxford University Press), 194–209.

Fuchs, T. (2017). "Levels of empathy–primary, extended, and reiterated empathy," in *Empathy*, eds V. Lux and S. Weigel (Berlin: Springer), 27–47. doi: 10.1057/978-1-137-51299-4_2

Fuchs, T., and De Jaegher, H. (2009). Enactive intersubjectivity: participatory sense-making and mutual incorporation. *Phenomenol. Cogn. Sci.* 8, 465–486. doi: 10.1007/s11097-009-9136-4

Fuchs, T., and Koch, S. C. (2014). Embodied affectivity: on moving and being moved. *Front. Psychol.* 5:508. doi: 10.3389/fpsyg.2014.00508

Gallagher, S. (2004). Understanding interpersonal problems in autism: interaction theory as an alternative to theory of mind. *Philos. Psychiatry Psychol.* 11, 199–217. doi: 10.1353/ppp.2004.0063

Gallagher, S. (2005). *How the Body Shapes the Mind*. Oxford: Oxford University Press, doi: 10.1093/0199271941.001.0001

Gallagher, S. (2012a). Empathy, simulation, and narrative. *Sci. Context* 25, 355–381. doi: 10.1017/s0269889712000117

Gallagher, S. (2012b). "Neurons, neonates and narrative," in *Moving Ourselves, Moving Others*, eds Ad Foolen, U. M. Lüdtke, T. P. Racine, and J. Zlatev (Amsterdam: Benjamins), 167–196.

Gallese, V. (2003). The roots of empathy: the shared manifold hypothesis and the neural basis of intersubjectivity. *Psychopathology* 36, 171–180. doi: 10.1159/000072786

Gallese, V., and Sinigaglia, C. (2018). "Embodied resonance," in *The Oxford Handbook of 4E Cognition*, eds A. Newen, L. De Bruin, and S. Gallagher (Oxford: Oxford University Press), doi: 10.1093/oxfordhb/9780198735410.013.22

Gangopadhyay, N. (2014). Introduction: embodiment and empathy, current debates in social cognition. *Topoi* 33, 117–127. doi: 10.1007/s11245-013-9199-2

Gibson, J. J. (1979). *The Ecological Approach to Visual Perception: Classic Edition*. East Sussex: Psychology Press.

Gill, C. (2006). *The Structured Self in Hellenistic and Roman Thought*. Oxford: Oxford University Press.

Goldman, A. (2006). *Simulating Minds The Philosophy, Psychology, and Neuroscience of Mindreading*. New York, NY: Oxford University Press, doi: 10.1093/0195138929.001.0001

Goldman, A., and de Vignemont, F. (2009). Is social cognition embodied? *Trends Cogn. Sci.* 13, 154–159.

Gopnik, A., and Wellman, H. M. (1992). Why the child's theory of mind really is a theory. *Mind Lang.* 7, 145–171. doi: 10.1111/j.1468-0017.1992.tb00202.x

Graver, M. R. (2007). *Stoicism & Emotion*. Chicago, IL: University of Chicago Press.

Guterstam, A., Björnsdotter, M., Gentile, G., and Ehrsson, H. H. (2015). Posterior cingulate cortex integrates the senses of self-location and body ownership. *Curr. Biol.* 25, 1416–1425. doi: 10.1016/j.cub.2015.03.059

Hierocles (2009). *Hierocles the Stoic: Elements of Ethics, Fragments and Excerpts*. trans. I. Ramelli and D. Konstan Atlanta, GA: Society of Biblical Lit.

Hutto, D. D. (2004). The limits of spectatorial folk psychology. *Mind Lang.* 19, 548–573. doi: 10.1111/j.0268-1064.2004.00272.x

Inwood, B. (1985). *Ethics and Human Action in Early Stoicism*. New York, NY: Oxford University Press.

Jensen, R. T., and Moran, D. (2012). Introduction: intersubjectivity and empathy. *Phenomenol. Cogn. Sci.* 11, 125–133. doi: 10.1007/s11097-012-9258-y

Klein, J. (2016). The stoic argument from Oikeisis. *Oxf. Stud. Anc. Philos.* 50, 143–200. doi: 10.1093/acprof:oso/9780198778226.003.0005

Laertius, D. (1925). *Lives of Eminent Philosophers, Volume II* Loeb Classical Library 185. trans. R. D. Hicks Cambridge, MA: Harvard University Press.

León, F., and Zahavi, D. (2016). "Phenomenology of experiential sharing: the contribution of schutz and walther," in *The Phenomenological Approach to Social Reality*, eds A. Salice and H. Schmid (cham: Springer), 219–234. doi: 10.1007/978-3-319-27692-2_10

Long, A. A. (1996). "Hierocles on oikeiōsis and self-perception," in *Stoic Studies*, ed. Long (Berkeley, CA: Cambridge University Press), 250–263.

Long, A. A., and Sedley, D. N. (1987). *The Hellenistic Philosophers*. Cambridge: Cambridge University Press, 1.

Lyons-Ruth, K., Bruschweiler-Stern, N., Harrison, A. M., Morgan, A. C., Nahum, J. P., Sander, L., et al. (1998). Implicit relational knowing: its role in development and psychoanalytic treatment. *Infant Ment. Health J.* 19, 282–289. doi: 10.1002/(sici)1097-0355(199823)19:3<282::aid-imhj3>3.0.co;2-o

Maister, L., Sebanz, N., Knoblich, G., and Tsakiris, M. (2013). Experiencing Ownership over a dark-skinned body reduces implicit racial bias. *Cognition* 128, 170–178. doi: 10.1016/j.cognition.2013.04.002

Maselli, A., and Slater, M. (2013). The building blocks of the full body ownership illusion. *Front. Hum. Neurosci.* 7:83. doi: 10.3389/fnhum.2013.00083

Merleau-Ponty, M. (1962). *Phenomenology of Perception*. Abingdon: Routledge.

Newen, A., De Bruin, L., and Gallagher, S. (2018). *The Oxford Handbook of 4E Cognition*. Oxford: Oxford University Press.

O'Shea, J. R. (2012). The 'theory theory'of mind and the aims of sellars' original myth of jones. *Phenomenol. Cogn. Sci.* 11, 175–204. doi: 10.1007/s11097-011-9250-y

Peck, T. C., Seinfeld, S., Aglioti, S. M., and Slater, M. (2013). Putting yourself in the skin of a black avatar reduces implicit racial bias. *Conscious. Cogn.* 22, 779–787. doi: 10.1016/j.concog.2013.04.016

Pembroke, S. G. (1971). "Oikeiosis," in *Problems in Stoicism*, ed. A. A. Long (Oxford: Athlone Press), 126.

Petkova, V. I., Björnsdotter, M., Gentile, G., Jonsson, T., Li, T., and Ehrsson, H. H. (2011). From part- to whole-body ownership in the multisensory brain. *Curr. Biol.* 21, 1118–1122. doi: 10.1016/j.cub.2011.05.022

Plato (1997). *Complete Works*, eds J. M. Cooper and D. S. Hutchinson (Indianapolis, IN: Hackett Publishing).

Plutarch (1976). *Moralia, Volume XIII: Part 2: Stoic Essays* Loeb Classical Library 470, ed. J. Henderson (Cambridge, MA: Harvard University Press). trans. H. Cherniss.

Premack, D., and Woodruff, G. (1978). Does the chimpanzee have a theory of mind? *Behav. Brain Sci.* 1, 515–526. doi: 10.1017/s0140525x00076512

Scheler, M. (1954). *The Nature of Sympathy*. trans. P. Heath Abingdon: Routledge.

Schmidsberger, F., and Löffler-Stastka, H. (2018). Empathy is proprioceptive: the bodily fundament of empathy–a philosophical contribution to medical education. *BMC Med. Educ.* 18:69. doi: 10.1186/s12909-018-1161-y

Schutz, A. (1967). *The Phenomenology of the Social World*. Evanston, IL: Northwestern University Press.

Seneca, L. A. (2007). *Seneca: Selected Philosophical Letters - Translated with Introduction and Commentary*. trans. B. Inwood Oxford: University Press, Incorporated.

Slaby, J., and Stephan, A. (2008). Affective intentionality and self-consciousness. *Conscious. Cogn. Soc. Cogn. Emot. Self Conscious.* 17, 506–513. doi: 10.1016/j.concog.2008.03.007

Sorabji, R. (1993). *Animal Minds and Human Morals: The Origins of the Western Debate*. Ithaca, NY: Cornell University Press.

Stueber, K. (2006). *Rediscovering Empathy: Agency, Folk Psychology, and the Human Sciences*. Cambridge, MA: Institute of Technology.

Zahavi, D. (2008). Simulation, projection and empathy. *Conscious. Cogn. Soc. Cogn. Emot. Self Conscious.* 17, 514–522. doi: 10.1016/j.concog.2008.03.010

Zahavi, D. (2010). Empathy, embodiment and interpersonal understanding: from lipps to schutz. *Inquiry* 53, 285–306. doi: 10.1080/00201741003784663

Zahavi, D. (2011). Empathy and direct social perception: a phenomenological proposal. *Rev. Philos. Psychol.* 2, 541. doi: 10.1007/s13164-011-0070-3

Zahavi, D. (2014). Empathy and other-directed intentionality. *Topoi* 33, 129–142. doi: 10.1007/s11245-013-9197-4

Zahavi, D., and Michael, J. (2018). "Beyond mirroring: 4E perspectives on empathy," in *The Oxford Handbook of 4E Cognition*, eds A. Newen, L. De Bruin, and S. Gallagher (Oxford: Oxford University Press), 589–606.

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

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

CONCEPTUAL ANALYSIS
published: 21 April 2021
doi: 10.3389/fpsyg.2021.530560

Why Build a Robot With Artificial Consciousness? How to Begin? A Cross-Disciplinary Dialogue on the Design and Implementation of a Synthetic Model of Consciousness

David Harris Smith[1*†] *and Guido Schillaci*[2,3†]

[1] *Communication Studies and Media Arts, McMaster University, Hamilton, ON, Canada,* [2] *Department of Excellence in Robotics & AI, Scuola Superiore Sant'Anna, Pisa, Italy,* [3] *The BioRobotics Institute, Scuola Superiore Sant'Anna, Pisa, Italy*

Creativity is intrinsic to Humanities and STEM disciplines. In the activities of artists and engineers, for example, an attempt is made to bring something new into the world through counterfactual thinking. However, creativity in these disciplines is distinguished by differences in motivations and constraints. For example, engineers typically direct their creativity toward building solutions to practical problems, whereas the outcomes of artistic creativity, which are largely useless to practical purposes, aspire to enrich the world aesthetically and conceptually. In this essay, an artist (DHS) and a roboticist (GS) engage in a cross-disciplinary conceptual analysis of the creative problem of artificial consciousness in a robot, expressing the counterfactual thinking necessitated by the problem, as well as disciplinary differences in motivations, constraints, and applications. We especially deal with the question of why one would build an artificial consciousness and we consider how an illusionist theory of consciousness alters prominent ethical debates on synthetic consciousness. We discuss theories of consciousness and their applicability to synthetic consciousness. We discuss practical approaches to implementing artificial consciousness in a robot and conclude by considering the role of creativity in the project of developing an artificial consciousness.

Keywords: artificial consciousness, synthetic consciousness, robotics, art, interdisciplinary dialogue, synthetic phenomenology

OPEN ACCESS

Edited by:
Chiara Fini,
Sapienza University of Rome, Italy

Reviewed by:
Alfredo Paternoster,
University of Bergamo, Italy
Aaron Kozbelt,
Brooklyn College and the Graduate Center of the City University of New York, United States

****Correspondence:***
David Harris Smith
dhsmith@mcmaster.ca

†These authors have contributed equally to this work

Specialty section:
This article was submitted to Theoretical and Philosophical Psychology, a section of the journal Frontiers in Psychology

Received: *29 January 2020*
Accepted: *12 March 2021*
Published: *21 April 2021*

Citation:
Smith DH and Schillaci G (2021) Why Build a Robot With Artificial Consciousness? How to Begin? A Cross-Disciplinary Dialogue on the Design and Implementation of a Synthetic Model of Consciousness. Front. Psychol. 12:530560. doi: 10.3389/fpsyg.2021.530560

1. WHY BUILD AN ARTIFICIAL CONSCIOUSNESS?

1.1. DHS

Human culture owes much to the wish to animate matter, since we are largely constituted in our abilities and status in the world by investing the world with anthropomorphic meaning and agency far into our prehistory (Mithen and Morton, 1996). From stone, bone, and pigments, to writing and print, to sound, image and cinema, to artificial agents, one can trace a progressive anthropomorphic investment in our symbolic technologies, which are now capable of materializing and automating our imaginations, our words, our stories, our storytellers, our conviviality, and our intelligence. It is difficult to imagine that this trajectory will suddenly be arrested. Given the centrality of innovative anthropomorphism to cultural progress, the technical investment of consciousness appears inevitable. But, to what end? What artistic uses can be made of artificial consciousness, especially in the context of robots?

To consider this question, there is an important distinction to be made between the actual realization of sentience in robots and the works, stories, and myths, about sentient robots. There are numerous examples of the latter dating at least from the myth of Talos (700 B.C.) to contemporary films, such as Alex Gardner's Ex-Machina (Gardner, 2014). Wherever, robot artworks have been physically created with phenomenological premises, traits associated with consciousness, such as self-awareness, intention and emotion, are simulated rather than realized, for example Robot K-456 (1964) by Nam June Paik and Shuya Abe (Kac, 1997), Helpless Robot (1987) by Norman White (White, 1987), and hitchBOT (Smith and Zeller, 2017). I would propose that these works and stories about sentient robots stem from contemplation of the limits of human technological agency and the hazards of transgressing what has been "designed" by nature. To embark upon the project of building a robot with artificial consciousness would convert our speculations and apprehensions into design problems and inaugurate an entirely new domain in the arts concerned with the production of autonomous creativity and the deliberate craft of human-AI culture.

One promising direction for this project is Gallese's (2017) bio-cultural approach to art and aesthetics, which is grounded in embodied cognitive processes.

> The body literally stages subjectivity by means of a series of postures, feelings, expressions, and behaviors. At the same time, the body projects itself in the world and makes it its own stage where corporeality is actor and beholder; its expressive content is subjectively experienced and recognized in others (p. 181).

The material presence of a technical implementation of consciousness, allows us to confront the physical constitution of sentience. The presence of such a lively thing, as something that must be engaged spatially and socially, automates a tacit understanding of the physical constitution of our own experience of consciousness. There are, of course, other pathways to understanding the physical nature of human consciousness, or the illusion of consciousness, for example through scientific explanation but, for art, it is the presence of the aesthetic object that convenes experience and understanding.

1.2. GS

As a cognitive roboticist, I try to make machines come alive. Consciousness is one the most profound aspects that characterize us as human beings. Whether and how conscious machines that are aware of themselves can be created is an actively debated topic also in the robotics and artificial intelligence communities. Consciousness and self-awareness are, however, ambiguous terms and numerous theories about what constitute them have been proposed. A phenomenological account of consciousness has recently re-gained vigor in philosophy and brain sciences, which focuses on a low-level, pre-reflective aspect of consciousness: the *minimal self* (Gallagher, 2000; Metzinger, 2020). *Pre-reflective* stands for something that is experienced *before* rationally thinking about it, and mainly relates to the perception of our own body and the feeling of being in control of our own movements. This aspect of consciousness is perhaps the most easily accessible in terms of experimental exploration and quantification, and a number of measures and behavioral paradigms have been proposed in the literature (see Georgie et al., 2019 for a review). Empirical research supports the idea that such low-level subjective experiences rely on self-monitoring mechanisms and on predictive processes implemented by our brains.

Robots share similar characteristics with animals and humans: they are embodied agents that can act and perceive what is happening around them. Complex behaviors and internal representations can emerge from the interaction between their embodiments, the environments they are situated in, and the computational models implemented in their embedded computers. Building, monitoring, and analysing them, may provide insights in the understanding of different aspects of cognition, of subjective experiences (Schillaci et al., 2016; Lang et al., 2018), and of consciousness (Holland and Goodman, 2003; Chella et al., 2019).

2. WHAT ARE THE ETHICAL ISSUES?

2.1. DHS

The viability of artificial consciousness is often conceived as dependent upon the development of artificial general intelligence (AGI), consciousness regarded as an emergent property of general intelligence. Although this is certainly the case in the evolution of consciousness in human beings, there is no reason to suppose that consciousness will come along for the ride in the development of AGI. The association between general intelligence and consciousness also leads some to assume that artificial consciousness has similar development challenges. This may not be the case, and we won't know without separating the project of artificial consciousness from the project of artificial general intelligence.

The model of consciousness one proposes to implement affects the formulation of an ethics of artificial consciousness. The contingent mapping of ethics to models of consciousness can be organized around the themes of suffering, moral obligation, and alignment.

A proposal distinguishing suffering from pain in human beings, regards suffering as a type of avoidance or resistance to the experience of pain, which perversely amplifies and prolongs the painful experience. Suffering in this view implicates self-knowledge and the role of language in reflecting upon and abstracting experience.

> A specific process is posited as the source of the ubiquity of human suffering: the bidirectionality of human language. Pain is unavoidable for all complex living creatures, due to the exigencies of living, but human beings enormously amplify their own pain through language. Because verbal relations are arbitrarily applicable, any situation can "remind" humans of past hurts of all kinds. In nonverbal organisms, only formally similar situations will perform this function (Hayes, 2002, p. 62).

The problem of suffering in artificial consciousness as described by Metzinger (2018) derives from the assumption that consciousness and, in particular, a phenomenal self model, underwrites the capacity for suffering in human beings. Metzinger reasons that an artificial consciousness possessing human-like phenomenological models will have the potential to suffer as a result of poor or malicious design, thus it would be immoral to create an artificial consciousness. And worse, our copy-paste technologies would allow unlimited multiplication of suffering artificial patients. This is a challenging argument and one that yields some interesting questions when explored. For example, arguments for the avoidance of suffering are not reserved to the artificial. Is it not also an anti-natalist recommendation against human procreation? We should not have children by this account. Such arguments run contrary to the optimistic disposition of the majority of humankind and downplay our capacity for creative problem-solving in the face of novelty.

The moral implications of the claim that consciousness entails suffering varies depending upon whether this is a correlational or causal claim and how we think of suffering in relation to existential threats and physical damage. If what we call suffering is how the brain represents perceived threats and actual damage to our bodies, then consciousness is merely correlated with suffering. The absence of consciousness does not remove actual threats, and nor does it obviate real damage to human or animal bodies. However, the psychological nature of suffering appears to exceed this reductive correlation, particularly the types of suffering associated with remembrance, attention, and anticipation.

Psychological suffering entails attending to mental representations of pain, deprivation, revulsion, grief, anxiety, fear, and shame. Here one finds an interesting overlap between the role of mental representations and the claim that suffering might result from poorly designed artificial consciousness. Is not human suffering also a design problem entailing responses to accidents and surprises, the behavioral decisions of self and others, and our cognitive habits of representation? For example, Buddhist contemplative practices, while not claiming to resolve the causes of suffering from actual threats and real damage that come with our physical mortality, do attempt to re-frame the psychological experience of suffering through compassion, observation, and attentional training (Yates et al., 2017). Why would we not include criteria for psychological framing in the design of artificial consciousness? Metzinger does propose an applied ethics for the limited design and development of consciousness technologies, hinging upon the question, "What is a good state of consciousness?" (Metzinger, 2009).

Bryson's argument against moral obligation to machines (Bryson, 2016) also responds to the problem of multiplication of artificial patients. Bryson is pragmatic about the scope and scale of problems confronting humanity and our limited capacity to reserve care and resources to the needs of humans and animals, rather than robots, in the present and near future. Bryson does, however, consider that artificial consciousness may have creative application within the arts (Bryson et al., 2017). The latitude for experimentation with artificial consciousness within the arts may be justified by the voluntary participation of arts audiences in low-risk settings where fictions are expected.

To the extent that consciousness or, at least, the user-illusion of consciousness and self, have come to be associated with autonomy, Dennett (2019) argues that these features, in the absence of human vulnerability and mortality, would render an artificial consciousness indifferent to human values. The technical immortality of the artificial consciousness, its copy-paste methods for reproduction, and its on-off-and-on-again resistance to "death," certainly divide the machine bearers of consciousness from the human bearers, according to susceptibility to threat and damage. But this difference does not necessitate misalignment. It is possible that the resilience of artificial consciousness in the face of existential threats has something to teach us about the design of our own experiences in the context of mortality. For example, an effectively immortal artificial consciousness may not be subject to the limits of imagination associated with our lifespan horizons, for example by engaging in counterfactual thinking conducive to the welfare of multiple generations of humanity into the future.

2.2. GS

Implementing conscious machines would raise, indeed, different ethical concerns. Should they be considered as objects or as living agents? Studies have shown that simple social cues already strongly affect our views of robots. For instance, people refuse to turn off a small humanoid robot when it is begging for its life (Horstmann et al., 2018), or feel the destruction of a robot—as your *hitchBOT* taught us—morally wrong (Smith and Zeller, 2017; Fraser et al., 2019).

Should conscious machines have moral competence? Making moral decisions may require empathy with pain, suffering and emotional states of others (Wallach et al., 2011). Is building conscious robots that undergo pain and suffering ethical itself? As you pointed out, the moral implications of creating suffering artificial agents, as well as of claiming that consciousness entails suffering, may vary also depending on whether we think of suffering as a mere physical damage or as a higher mental representation of experiences of negative valence, perhaps over a longer time scale.

How can we assess whether robots could go through pain and suffering, though? Even the detection and assessment of pain in animals and insects is problematic. Animal scientists have been trying to define concepts and features that can be used to evaluate the potential for pain in vertebrates and invertebrates—to name a few: the possession of nociceptors, the existence of neural pathways from nociceptors to the brain, the capability to avoid potentially painful stimuli through learning, and so the like (Sneddon et al., 2014).

Recent accounts propose that the experience of pain, as well as subjective and emotional experience, results from a perceptual inference process (Seth et al., 2012; Pezzulo, 2017; Kiverstein et al., 2019). This would explain, for instance, how pain perception seems to be affected not just by physical damages but also by past experiences, expectations and emotions (Garcia-Larrea and Bastuji, 2018). I believe that modeling these processes in robots—and integrating them within a

bigger framework where behaviors are driven by different types of imperatives and goals—may help in shedding light on the nature and valence of pain, suffering, and consciousness in humans.

3. WHAT IS REQUIRED FOR AN ARTIFICIAL CONSCIOUSNESS?

3.1. DHS

A naturalist theory of consciousness necessitates an evolutionary explanation of how simple organisms could evolve complex minds capable of the type of intelligent and reflexive cognitive features we associate with subjective experience. One type of evolutionary explanation proposes that consciousness arises spontaneously given some sufficient degree of complexity and integration in the information processing capacity of a biological, or indeed, a physical or technical system (for example, see Tononi, 2012). Proposing an informational approach that is tightly bound to biological life, Damasio (2012) considers the adaptive advantages of a successive stages of evolving self-modeling processes: the protoself representing vital information or primordial feelings about the body and status of the organism, the core self representing information about its interactions with other organisms, objects, and environments, and the autobiographical self comprised by complex representations combining core self and protoself with memory and future simulation. Features of consciousness associated with the autobiographical self have evolved, perhaps uniquely, in humans coincident with language and culture: "Consciousness in the fullest sense of the term emerged after such knowledge was categorized, symbolized in varied forms (including recursive language), and manipulated by imagination and reason" (Damasio, 2012, p. 182). An information-based theory of consciousness would need to process, integrate, and resolve low level incoming information with these higher-order predictive representations. Ultimately we would look to neuroscience for plausible mechanisms and implementations that integrate bottom-up and top-down information, for example Dendritic Information Theory (Aru et al., 2020).

While all naturalist theories of consciousness are equal in their status as provisional, rather than generally accepted scientific explanations, the pragmatic aim of building a synthetic consciousness recommends against the most speculative of these theories at this time, including quantum theories of consciousness (Hameroff and Penrose, 1996) and panpsychist assertions that consciousness is a fundamental (yet currently undetected) physical feature of the universe (Goff et al., 2001). I am suspicious of theories of consciousness, hijacking the anthropomorphic principle, that begin with the assertion that since we live in a universe where consciousness exists, it must therefore be a fundamental feature of the universe. Imagine replacing "consciousness" with "duck-down duvets" and you will see the troubles piling on.

This leaves in place a candidate group of information theories of consciousness that attempt to model brain-based biophysical information processes in a variety of framings, including lower level theories, which ground explanations in neural processes, and higher order theories emphasizing mental representations. A naturalistic account of consciousness maintains that phenomenal consciousness is an effect, or result, of brain functions and mental representations. These can be accounted for in higher-order cognitive theories that explain consciousness in terms of causal role, having a function in an architecture of mental processes and intentional contents. Mental states that are considered to be phenomenal consciousness "are those states *that possess fine-grained intentional contents of which the subject is aware*, being the target or potential target of some sort of higher-order representation" (Carruthers, 2016).

Thagard (2019) employs a "three-analysis" using exemplars, typical features, and explanations, to approximate a pragmatic definition of consciousness. What are typical, or broadly accepted examples of consciousness, what features do we associate with consciousness, and how is consciousness used in explaining other phenomena? Exemplars of consciousness are sensory perceptions and perceptions of pain, emotions, thoughts, and self-awareness. Typical features of consciousness include experience, attention, wakefulness, and awareness. Consciousness figures in explanations of voluntary behavior, self-reports, and wakefulness (Thagard, 2019, p. 159–160). To complete a list of ingredients for consciousness that we could use as a design specification for an artificial consciousness, I would add features identified by Metzinger (2009), such as an integrated self and world model that is continuously updating and some kind of temporal icon to provide a locus of first-person perspective in the flow of experience over time—a *now*.

The question "What causes us to report having conscious experiences?" sets aside any substantive claims about consciousness as some special kind of "stuff." This is the research question proposed by Graziano (2016, 2019) and one which is broadly consistent with information-based illusionistic theories of consciousness (Dennett, 1991, 2016; Frankish, 2016): "To understand consciousness, we need to look for a system in the brain that computes *information about consciousness*—about its properties and consequences" (Graziano, 2019, p. 77–78). I assume consciousness to be a subset of the total of cognitive processes of the brain and body and find it plausible that the experience of consciousness consists of a reductive, and likely predictive, representation of the brain's attentional activities and intentional contents, or an attention schema (Graziano, 2018; Graziano et al., 2020). Here, it is important to highlight controversies about the nature of attention, in particular the attempted distinctions between attention, intention, and awareness, which might be more usefully subsumed under the concept of cognitive "selection" (Hommel et al., 2019).

The attention schema might also serve as a temporal icon, providing an ongoing, stable sense of presence, or "now," in the brain's continuous updating of sequential selections. The representation of a "now" would rely upon event driven processes to mark time. The sources of events in body/brain system are attentional shifts stimulated by either mindwandering or environmental inputs, or possibly interoception of the autonomous rhythms of heartbeats and respiration. Regardless of source, an abstract representation of event driven perceptions

would form the contents of type of fleeting *memory of the present* from which a sense of the immediate present, or "now" is abstracted (see also fragile short term memory in Block, 2007, 2011). In this configuration, short term memory provides a gestalt representation of the now; it feels rich, but in much the same way that a visual scene appears to be rich and complete in its detail despite its fragmentary construction by the visual system.

In fact, gestalt effects typical of visual perception, seem to be a good analogy for the phenomenology of consciousness, its feel of ineffable wholeness and ubiquity arising from piecemeal cognitive processes giving the predictive illusions of closure, similarity, and continuity. Assuming that consciousness is a reductive subset of the total of the brain's cognitive processes, a naive feature of cognitive impenetrability is required for consciousness to maintain and utilize a model of a durable observing self that believes it has global and holistic access to, and possession of, the moment-to-moment contents of experience. This naiveté is central to being a subject of conscious experience (Metzinger, 2009; Graziano, 2018; Graziano et al., 2020).

I have assembled the following table of proposed variables contributing to the phenomenology of consciousness from the ideas and literature cited above. These can be variables can serve as design criteria for an artificial consciousness. I have simplified, in some cases, by collapsing several variables under one label.

This list of variables in **Table 1** could be used as a guide to the features of an artificial consciousness in a robot.

TABLE 1 | List of variables contributing to reports of conscious experience.

Variable	Description
Body	A physical implementation with optimal duration or homeostasis. Since we are modeling a naturalist explanation of conscious experience, a body or physical implementation is required. Information is substrate independent, nevertheless, it requires a physical form to do something. Homeostasis is added to provide a needed value to animate the body and to distinguish salient information.
Wakefulness	Variable states of responsiveness or arousal, for example: from comatose, to dreaming, to vigilance. A minimal level of responsiveness is a pre-condition for having conscious experience.
Action	Capacity to cause changes in physical domain, including cognitive domain (information, while substrate independent requires physical implementation).
Perception	Mechanisms for sensing and representing physical domain, including cognitive domain.
Searchable memory	Mechanism and processes for short and long term retention and retrieval of representations.
Integrated self and environment model	Updatable reductive, abstract representations of "I" and "me," "my body," character, personality, narrative, and counterfactual self. Updatable reductive, abstract representations of physical body, others, environment, physics, and the arrow of time.
Integrated attention, intention, and temporal schema	Updatable reductive, abstract representation of perceptual attention, and intentional status. An iconic representation marking the present moment in a sequential flow of events, providing an updatable locus of perspective vis-a-vis intentional representations.
Language	Semantic and linguistic representation to communicate reports of conscious experience.

3.2. GS

I tend to focus on low-level phenomenological aspects of consciousness. Contemporary phenomenologists (Zahavi and Parnas, 1998, 2003; Gallagher, 2006) argue that the most basic level of self-consciousness is the *minimal self*, i.e., "the pre-reflexive point of origin for action, experience, and thought" (Gallagher, 2000). Some scholars (see Zahavi) claim that the minimal self precedes any social dimension of selfhood, while others (Higgins, 2020) see this minimal form of experiential selfhood in humans as equiprimordial with socially constituted experiences. Primitive forms of sense of self developed in early infancy have been proposed to crucially rely on caregiver-infants close embodied relationship (Ciaunica and Fotopoulou, 2017; Ciaunica and Crucianelli, 2019), which allow the developing organism to further mentalize its homeostatic regulation of interoceptive signals (Fotopoulou and Tsakiris, 2017).

Higher-order theories of consciousness explain subjective experience throughout the cognitive ability of being aware of one's own mental states (see Lyyra, 2010 for an interesting review). Whereas higher-order theories of consciousness can be useful in differentiating forms of self-awareness, they do not offer a clear account of how it bootstraps and of how "infants or animals can undergo phenomenal experience without being aware of such phenomenal states" (Lyyra, 2010). I think that a more pragmatic approach to the implementation of a *developing* artificial consciousness would better start from more minimal forms of experiential selfhood, addressing low-level phenomenological aspects of consciousness.

Developmental psychologists and brain scientists have been seeking links between cognitive development and the experience of the minimal self. Studies showed that newborns are systematic and deliberate in exploring their own body and the consequences of their own actions, suggesting the gradual formation of causal models in their brains (Rochat, 1998; Rochat and Striano, 2000). Motor knowledge and proto-representations of the body seem to be forming already during pre-natal developmental stages (Zoia et al., 2007). Paradigms for measuring body awareness and agency attribution in infants (Filippetti et al., 2014; Filippetti and Tsakiris, 2018), as well as in adults (Shergill et al., 2003; Ehrsson et al., 2004), can be also found in the literature. As mentioned above, caregiver-infants close embodied relationship seems to support the development of primitive forms of a sense of self (Ciaunica and Fotopoulou, 2017; Ciaunica and Crucianelli, 2019).

These studies indicate emergent conscious phenomenology already during early developmental stages. But what is driving this process? What are the computational and behavioral prerequisites that would let this emerge also in robots? If we

take a developmental standpoint, some of the variables that you suggested in **Table 1** may be appearing at later stages of development, and others may be more intertwined. For instance, language may be not essential in early developmental stages of consciousness. Developmental psychologists measure subjective experience in infants through non-verbal indicators, e.g., looking time to visual stimuli, hemodynamic response measured through brain imaging techniques, number of movement units of their limbs, etc. An integrated self-representation seems to emerge throughout embodied interactions.

Experience affects perception, as well: what our brain perceives seems to be shaped by prior beliefs and expectations, according to the predictive brain hypothesis (Clark, 2013). The Free Energy Principle (FEP) (Friston, 2009, 2010) brings this forward, suggesting that brain functioning can be explained under the single imperative of minimizing prediction error, i.e., the difference between expected (or predicted) and perceived sensations (Pezzulo, 2017). Recent research posed a link between predictive processes, curiosity and learning (Oudeyer et al., 2007), and emotional experience (Kiverstein et al., 2019). According to these proposals, biological systems not only track the constantly fluctuating instantaneous errors, but also pay attention to the dynamics of error reduction over longer time scales. Interacting with the environment as part of epistemic foraging may generate more prediction error, but nonetheless may feel good for the agent. I find these studies extremely interesting and I feel that these processes may have a role also in conscious experience. Analysing the rate at which those errors are being reduced or increasing over time may provide insights about emotional engagement in humans and its implementation in artificial system. In a recent study with Alejandra Ciria and Bruno Lara, we showed that linking prediction error dynamics, emotional valence of action and self-regulatory mechanisms can promote learning in a robot (Schillaci et al., 2020a). The generative models that realize adaptive behaviors in biological systems may be driven by different drives (Pezzulo, 2017). Self-regulatory mechanisms should be also taken into account in the development of an artificial consciousness.

3.3. Complementary Strategies

In summary, two complementary approaches to the challenge of building an artificial consciousness are taken here. DHS tends toward a higher-order theory of consciousness, focusing on the importance of mental representations, such as primordial to complex self models and their contribution to conscious phenomenology. GS takes a lower-level approach which seeks to explain phenomenal, minimal self-experiences by means of embodied and computational processes, such as predictive processes. He presumes that embodied interactions with the world and with other individuals support the gradual formation of internal models and representations, ultimately allowing reflective conscious phenomenology at later stages of the developmental process.

Both DHS and GS converge on naturalist, developmental and brain-based explanations of the evolution and emergence of conscious experience.

4. HOW TO BEGIN?

4.1. DHS

Assuming a higher order theory of consciousness, the variables that contribute to conscious experience need to be modeled in an architecture of representations derived from fine-grained neural activity. How the brain's neural representations are encoded and related in such an architecture is an open question. As I understand it, approaches to encoding and decoding higher order representations can proceed by either attempting to imitate what the brain does when it construes complex representations, or by following computational methods that might achieve similar results by different means.

I am not sure where I first encountered the analogy (maybe Edwin Hutchins?), but I like to think of this choice of computational algorithmic vs. implementation level approaches as *fish vs. submarine*. If you want to design and build something that can swim underwater you could try to manufacture an artificial fish in all of its detail, or you could build a submarine. The analogy helps me think about the advantages and disadvantages of the two approaches for artificial consciousness. Building a fish will produce the desired result eventually but might also consist in wasted research and development effort in the reproduction of trivial, or irrelevant features, such as how to achieve the unique variation in the colored speckles of trout skin. On the other hand, building a submarine may result in overlooking critical fish features, such as the friction drag reduction of the scales on trout skin. Ideally, an artificial consciousness designer would avail of the function approximating approach of submarine (computational) design, while drawing inspiration from the salient features of fish (brain) design.

For describing the functions and integration of cognitive systems giving rise to conscious experience, the attention schema in **Figure 1** (Graziano, 2016, 2019; Graziano and Webb, 2018; Graziano et al., 2020) for building artificial consciousness looks like a good place to begin. Graziano and Webb (2018) propose a design sketch of the key features required to build artificial consciousness. These include a layered set of cognitive models beginning with (1) objective awareness of something, such as a perception of an apple, (2) cognitive access, or an information search and retrieval capability with a linguistic interface that can report information on the machine's internal models, (3) a self-model, or information about the machine's body, history, capabilities, and (4) an attention schema which integrates the layers of objective awareness and self-modeling information and is able to report this integrated relationship. The attention schema represents the machine's current allocation of computing and sensor resources to the contents of its objective awareness and the relation of these intentional contents to the self-model.

The attention schema layer is also where phenomenological features are implemented. For example, the sense of subjective awareness as something that feels internal and approximately spatially anchored to the self-model and the sense that the contents of awareness are something possessed by the self and available to be acted upon by the self. A machine with the proposed layered cognitive features of object awareness, cognitive

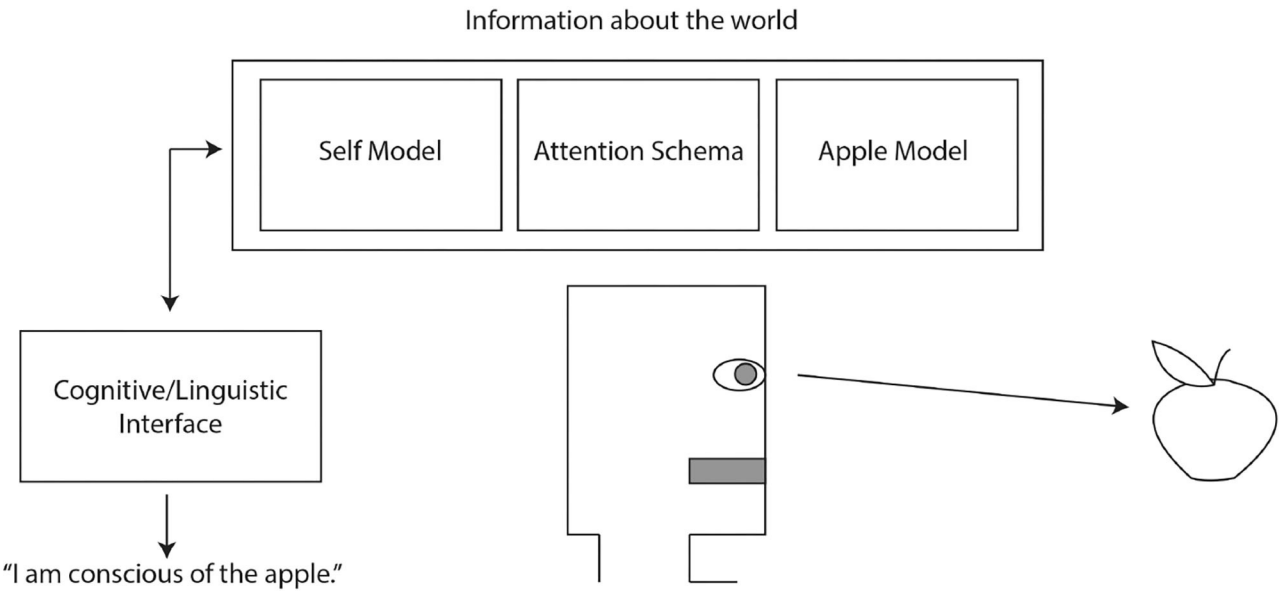

FIGURE 1 | Adapted from Graziano (2019, p. 174). The attention schema incorporating cognitive features of objective awareness, cognitive access, and self-model.

access, self-model, and attention schema, should be able to report, "I have mental possession of an apple" (Graziano and Webb, 2018, 294).

Most importantly, the attention schema is also naïve about its own construction. Because the schema is only able to report on information that it has access to, and it does not have access to information about its own coding and hardware functions, the schema is *transparent* to itself; it suffers from cognitive impenetrability. Such a conscious machine could have a parallel set of information processes that are able to objectively monitor and report on how the whole system is put together "how the representational models are constructed—under the hood" (see Holland and Goodman, 2003 for a discussion of this transparency). This would be a machine that has one system for naïve subjective awareness and another system for objective analysis, very much like the much maligned homunculus philosopher of mind ☺.

An artificial consciousness would also require some overarching objective to guide its values for information seeking and constructing salient representations. For example, the varieties of information that a human self-model abstracts, such as physical body, sense of agency, and social status, are finely tuned to prioritize genetic replication. Defining and declaring these orienting values for self modeling in an artificial consciousness involves design decisions with moral and ethical implications (Metzinger, 2009, 2018; Dennett, 2019), thus "survival and/or replication" might not be the wisest choice for arbitrarily assigned values to guide the behavior of our artificial consciousness. A more genteel and human-compatible objective for a robot with artificial consciousness might be "to learn and model knowledge about human consciousness" with some safeguards to ensure that the robot's information seeking behaviors are the result of voluntary human-robot interactions and decidedly passive and observational in execution. Such an objective would necessitate modeling the values that shape human consciousness, providing an overlapping domain of aligned objectives between sentient machines and human beings. Adding values by design suggests that we are engaged in building a hybrid symbolic and deep learning model, one that relies upon both assigned and learned values.

Given the gap that exists between the type of fine-grained unstructured data generated by the robot's sensors and the complex representations required for an attention schema, we need a computational method for building complex representations. Semantic pointer architecture or SPA in **Figure 2** (Eliasmith, 2013; Thagard and Stewart, 2014; Thagard, 2019), models encoding of data into the type of layered cognitive models required in the attention schema. SPA models how multiple sources of granular information acquired in networks of lower level sensory and motor neurons can be formed into more complex representations, binding neural networks through pointers. SPA models how representations function by decomposition, or unpacking, to their constituent information networks and how neural network representations can point to or infer other complex representations. Competition among semantic pointers through recurrent connections among neurons provides a process which could support gestalt cognition, shifting attention, representing changes in experience, and mindwandering.

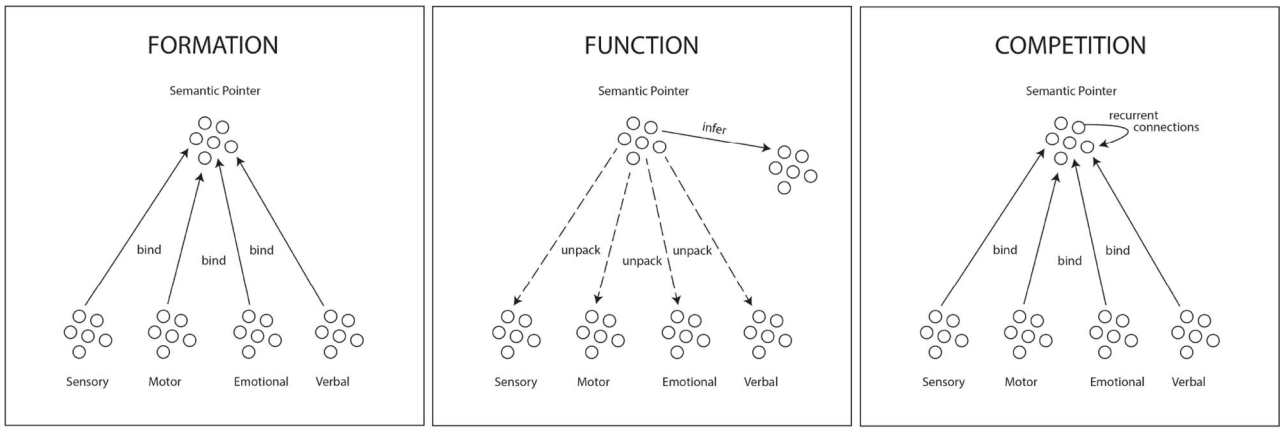

FIGURE 2 | Adapted from Thagard and Stewart (2014, p. 74–76). "Semantic pointers function to provide inferences by virtue of relations to other semantic pointers and can also unpack (decompress) into the sensory, motor, emotional, and/or verbal representations whose bindings formed the semantic pointer."

Assuming that we have, in the attention schema, a plausible theory of artificial consciousness, and a practical method for encoding and decoding neural networks to achieve its constituent cognitive models, what remains is to create an experimental design for the robot based on causal modeling and evidence testing.

A causal diagram (Pearl and Mackenzie, 2018) would indicate what causes the robot to have, and report, conscious experience. The diagram should incorporate the variables, or combinations of variables, listed in **Table 1**, all of which are explicit or assumed in the attention schema, as well as some type of intervention to activate the chain of cause and effect (see **Figure 3**). In this case the intervention is the question, "Are you conscious?," posed to the robot. This is, in all likelihood, spectacularly wrong-headed, but I am more than happy to start with "wrong" so that I can enrich my life by starting interesting arguments with friends.

Body, world, and memory are variable sources of intentional relations. The robot's attention may be directed toward information coming from its body, its environment, and its memory, which would include successive updating loop of self models (proto, core, and autobiographical). Attention information supplies objective awareness the substance of consciously accessible perceptions and interoceptions. All informational contents are bound together in a semantic pointer architecture, such that domain specific information, like the body model, is actually composed of inferences and predictions from its constituent neural networks in the architecture. The neural architecture supporting the attention schema contributes unpackable lower-level information from cognitive processes related to the body, the self, the world, objective awareness, memory, and attention. There is no binding problem in this model of consciousness because the attention schema is a gestalt-like prediction generated by this architecture. Memory and the informational contents of objective awareness inform the self and world models. The profile of objective awareness, which is constituted by a variable emphasis of the combined subjective and objective models, informs the attention schema. The schema, from a phenomenological perspective, is searchable because it is taken into short term memory and it may be queried and decomposed to its constituent world, or object, and self models. A short term memory loop may entail a type of buffering memory, with a fade-in prediction and fade-out memory gradient centered on an abstract representation of "now"—this would provide an always-advancing-into-the-future temporal icon upon which can be hung the "what it feels like" of conscious (hetero)phenomenology.

4.2. GS

Graziano's higher order theory of consciousness has some aspects that sound plausible to me, others rather more problematic. For instance, the proposal that conscious experience requires a model of the self, which would comprehend low-level bodily aspects and high-level autobiographical aspects of the self (Graziano and Webb, 2018), reminds me of Gallagher's distinction between minimal self and narrative self (Gallagher, 2000). As argued before, phenomenology of the self seems to emerge already during early infancy, likely before more complex, say autobiographical, models of the self develop.

Graziano also suggests that our brains maintain internal models of objects, and argues about the need of an objective awareness component: when sensory information about an object is available and is processed, the machine becomes objectively aware of that object. I subscribe to the idea that our brain makes up internal models of the world, but perception seems to have a more inferential, hypothesis testing nature than previously thought (Clark, 2013). This would already assign a subjective flavor to our awareness of the external world. Perception can be influenced by many other things, even by the presence or absence of action (see, for example, Troxler fading illusion reported in Parr et al., 2019 and in **Figure 4**).

Another comment is on the cognitive access component and the linguistic interface that—although not essential (Graziano and Webb, 2018)—would make the experimenter able to query

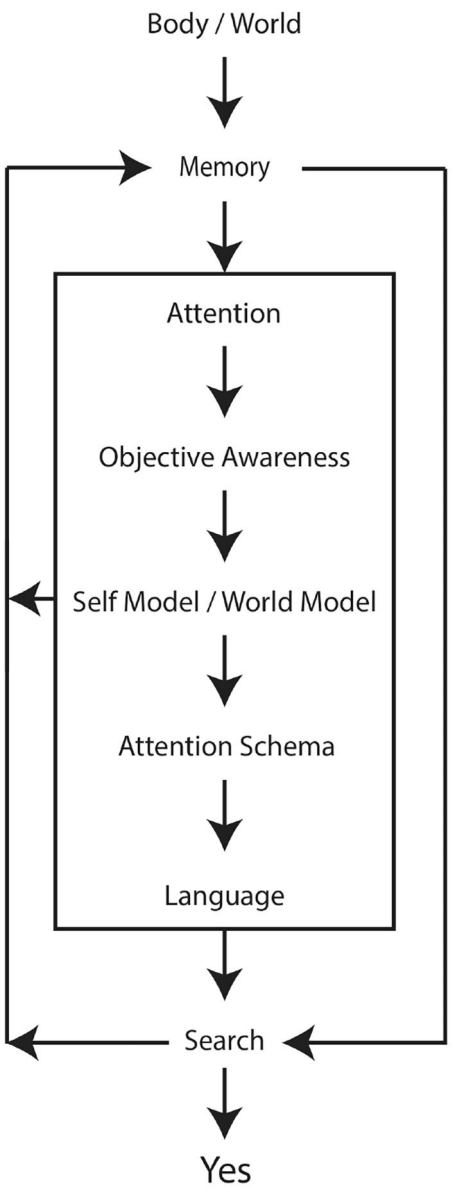

FIGURE 3 | Causal Diagram for Artificial Consciousness in a Robot. The arrows indicate direction of cause and effect. Reverse direction indicates "listens to," for example the self-model/world model listens to the objective awareness function, which in turn listens to the attention function. Attention, objective awareness, self and world models, the attention schema, and language also listen to memory and, in turn, shape memory. Downstream of body/world, all functions are proposed to be constituted by a searchable (unpackable) semantic pointer architecture.

the machine. However, different levels of consciousness can be attributed to animals and people from few behavioral features, without the need to engage in a conversation. I would thus explore—before the linguistic interface—which robot behaviors could induce us in the attribution of consciousness.

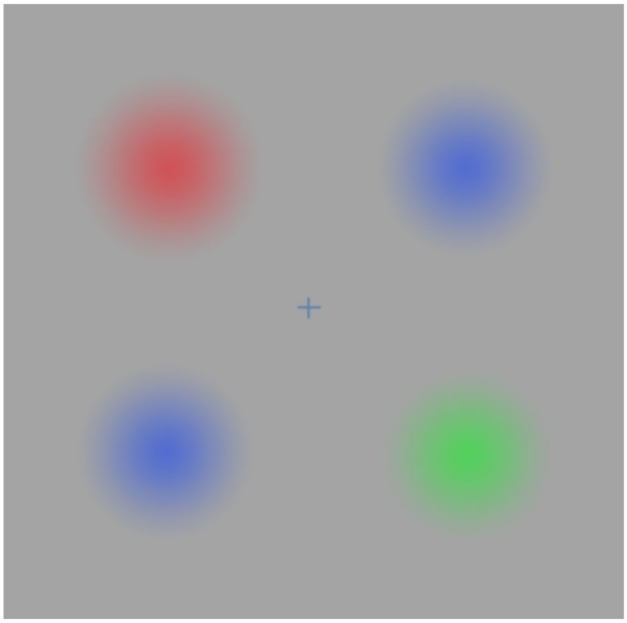

FIGURE 4 | Adapted from Parr et al. (2019). Troxler fading: when fixating the cross in the center of the image, the colors in the periphery gradually fade until they match the gray color in the background; when saccadic exploration is performed, colored blurred circles become visible.

I would also look at more robust methods to quantify subjective experience. In a recent paper, we discussed different paradigms and measures used in cognitive and brain sciences, and reviewed related robotics studies (Georgie et al., 2019). What would constitute a successful demonstration of artificial consciousness (Spatola and Urbanska, 2018)?

This also relates to the central element of Graziano's theory: the attention schema. Graziano suggests that the machine can claim it has subjective experience "because it is captive to the incomplete information in the internal models"—i.e., the models of the self and of the object, through an internal model of attention (Graziano and Webb, 2018). Subjective awareness of something would be thus "a caricature of attention." As he claims, if a machine can direct mechanistic attention to a specific signal, and if the machine has an internal model of that attentional state, then the machine can say that it is aware of that signal. I recognize that attentional processes may have an important role in conscious experience, as well as in perception and action, but this conclusion sounds too simplistic to me. Moreover, how would such an attention schema be concretely implemented? I find interesting an account that comes with the active inference proposal (Feldman and Friston, 2010), where attention is viewed as a selective sampling of sensory data that have high-precision in relation to the model's predictions. In a way, this is deeply intertwined with the agent's internal models, more than—as it sounds to me—as in Graziano's model. These comments would apply also to your causal diagram.

I find the semantic pointer architecture (SPA) interesting. Similar works on grounding complex representations on multi-modal experience can be found in the developmental robotics

literature. An example is the Epigenetic Robotics Architecture (ERA) (Morse et al., 2010), used for the acquisition of language and body representations in a humanoid robot. ERA self-organizes and integrates different modalities through experience. I have also studied similar models for the incremental learning of internal models (Escobar-Juárez et al., 2016; Schillaci et al., 2016), where representations were grounded on integrated motor and sensor maps, similarly to SPA. I also investigated how predictive capabilities could emerge from such representations, and how prediction errors could be exploited as cues for self-perception (Schillaci et al., 2016). Similar processes are thought to be involved in minimal self experiences.

4.3. DHS

As you point out, objective awareness entails prediction, but I think predictive processing is consistent with the attention schema model through input of salience values and prior conditioning and the role these play in perceptions associated with objective awareness. Additionally, objective awareness in AST is not exclusively reliant upon environmental inputs and gross physical actions, interoception and memory also supply inputs to objective awareness.

On the issue of verification of consciousness, admittedly the approach taken by AST of simply asking the robot if it is conscious seems facile (as I scurry off to read your papers). But I believe this superficial approach has merits that are specifically relevant to artificial consciousness and the AST model. In the AST model, human consciousness is an informationally impoverished representation of attention; representations of objects, the world, the self, and attention do not include information about the processes leading to representation in the brain. The self that claims to have conscious experience is ignorant of the neurological mechanisms that cause the claimed experience and experiencer. In this respect, it is an important evaluative tool to test for this ignorance. However, as evaluators of an artificial consciousness, we also have access to the systems of the AI that are impenetrable to itself. We can know and monitor the performance of the nested set of representations in our causal model, to see how they are engaged when the robot considers the query "Are you conscious?" In theory, we would have evaluative tools combining synthetic self-reports and quantitative measures of the systems producing the self-reports.

5. WHAT IS THE ROLE OF CREATIVITY IN ARTIFICIAL CONSCIOUSNESS?

5.1. DHS

The project of building an artificial consciousness engages with creativity in several contexts. First, there is the question of how synthetic consciousness will be included by artists in the materials and methods of art making. Much of contemporary art is motivated by politics, criticism, and reflexivity. While an art of artificial consciousness might become just another medium that artists may use to express these secular contents, its sentient aspirations might otherwise reinvigorate an aesthetics of existential wonder. Rather than promoting anthropocentric hubris, as some might claim, artificial consciousness confronts us with the humbling genesis of mind from matter, and the emergence of subjective experience in a non-differentiated physical field. In the case of a synthetic consciousness, our attention and critical appraisals must be directed to the form or medium of the artwork, rather than its ostensible contents. Often, in the discussion of consciousness, one encounters a division between the contents of consciousness and consciousness itself. Most artists will recognize a striking similarity between this distinction and the historic tensions between formalism (materials, methods, and ground) and representation (symbolism, reference, meaning) in art (Zangwill, 1999). The artistic engagement with artificial consciousness would constitute an unsurpassable formalism. After all, isn't consciousness the ground of all appearances and, ironically, itself an appearance?

Secondly, there is the creativity of the synthetic consciousness itself. An artificial consciousness will be an historic event in the human development and use of symbolic media, in this case, the technical investment of another kind of introspecting perspectival witness to the unfolding universe. Due to the transparent nature of its consciousness, this would be an artwork possessed of its own boredoms and uncertainties, and consequently prepared and motivated for the work of curiosity and creativity. Of course, creative functions leveraging uncertainty, such as mind-wandering behavior would require design and implementation. Mind-wandering requires the ability to combine representations in increasingly complex and novel formations and, importantly, to decompose representations to their constituent lower level representations. In this way, an artificial consciousness could travel the space of ideas, associating, assembling, disassembling, and reassembling unique proposals, in search of novel representations to satisfy its aesthetic values.

The cognitive scientist Margaret Boden describes three types of creativity: exploratory, combinatorial, and transformative (Boden, 2009). The first two types of creativity, exploratory and combinatorial, describe novel, or surprising outcomes as artists engage with either in-depth or hybrid investigations of familiar, rule-bound domains. Transformational creativity, on the other hand, constitutes "changing the rules," or a perturbation of these domains (Du Sautoy, 2020). Such perturbation according to Du Sautoy (2020), would likely stem from a disruption of our current assumptions about the role of free will in artistic creation,

> Our creativity is intimately bound up with our free will, something that it seems impossible to automate. To programme free will would be to contradict what free will means. Although, then again, we might end up asking whether our free will is an illusion which just masks the complexity of our underlying algorithmic process (Du Sautoy, 2020, p. 301).

The confrontation with artificial consciousness, with its phenomenological connotations of experience, creativity, and self-expression, might, as Du Sautoy suggests, motivate better explanations of the cognitive processes that appear to us as human creativity.

One of the projects of an artificial consciousness might be the discovery of unique aesthetic values, perhaps a sense of beauty that is salient only to the conscious machine. For example, in what ways would an artificial consciousness surprise us? Surprises of observational profundity, sensory pleasure, and narrative fulfillment, are what we have come to value in the arts, but I wonder what are the aesthetic possibilities of scientific creativity? Given the role of creativity in proposing scientific explanations and the knowledge that all scientific explanations are destined to be approximations of reality, is it possible that our artificial consciousness could use its transformative creativity to generate multiple novel, yet viable, approximations of reality, distinguished only by their aesthetics, their framing of the sublime? Science and art will converge in creative artificial consciousness.

5.2. GS

I agree with you that this project engages with creativity on many aspects: in the creative process of designing and building the artificial consciousness; in the new perspectives and possibilities that an artificial consciousness could open to artists; in developing conscious agents that are creative themselves.

We are not so far—I think—from having creative machines. There are examples out there of generative systems that can be used in explorative and creative processes—Google's deep dream, to name one, which is capable of generating novel visual artifacts from an initial knowledge of drawings and paintings. I believe that such systems would fit, however, within the category of "novel tools for creative people." They do broaden exploration possibilities, but the creativity of such algorithms is very much biased by their designer, who outlines the underlying AI machinery, decides how to train them and how they should explore, and eventually selects the best generated samples. Somehow, such AIs are given aesthetic values already from their creators.

I find very interesting your idea of studying whether and how aesthetic values could, instead, develop in a conscious learning machine. I can imagine that basic aesthetic values and drives could be given a priori by the designer. Then, I wonder whether this unique sense of beauty that you mention, which is salient only to the machine, could develop throughout its lifetime. Experiences may form attitudes and interests, shape the temperament and emotional engagement in the various activities, and consequently affect the aesthetic values and creativity of such an artificial agent.

The cognitive architecture you depicted can be in part implemented with tools that are currently under investigation in robotics and AI (see algorithms generating artificial curiosity and novelty-seeking behaviors; Schmidhuber, 2006; Oudeyer et al., 2007; Schillaci et al., 2020b). I think that the gap between curiosity and creativity, here, is small. Intrinsic motivation algorithms are driven by epistemic value "which correlates to the reduction of uncertainty" of an action, but could be designed also to be driven by aesthetic value. Would this be enough to produce a machine that develops a sense of beauty?

ALL TOGETHER NOW…

Although many of the issues featured in our dialogue are represented in the current literature, we hope that our discussion of the creative application of artificial consciousness helps to concretize these issues.

Consciousness appears to be a subset of the whole of human and animal cognitive activity, composed of composite and layered processes, rather than a singular process or yet-to-be-discovered substance. To design and build an artificial consciousness requires beginning with and resolving low-level processes which, further on, may develop complex higher order cognitive features, such as the autobiographical self. According to the reviewed proposals, the phenomenology of consciousness in human beings features a stream of selected representations that appear to be governed by competition in the context of limited cognitive resources and adaptive pressures for decisive action. This raises the possibility that consciousness is the result of constraints that are not necessarily the case in an artificial system with extensible computing capacity in low risk settings. Must we design artificial dangers and constraints in our artificial system to promote the phenomenology of a stream of consciousness, or rather, allow for multiple parallel streams of consciousness in a single entity?

We take seriously the ethical concern for the potential of artificial consciousness to suffer but we differ on the best course of action to take in response to this concern. It is within the realm of possibility that an artificial consciousness may happen by accident, for example in the case of a self-programming AI, and therefore we conclude that the deliberate project of designing an artificial consciousness capable of ameliorating its own suffering is an important undertaking and one which is at least the shared concern of the arts disciplines.

We have discussed low-level computational and behavioral features that we believe would be needed for building an artificial consciousness but admit to the difficulty of deriving the required higher order representations. We consider embodied interactions as of fundamental importance for the incremental learning of the dynamics of perceptual causality. It is upon embodied intentional experience and attentional capacity that, since early in life, we construct beliefs and expectations about ourselves, our bodies and our surroundings, and that we define values on internal and external goals. An artificial consciousness should employ computational mechanisms that allow such constructions. We consider creativity in all its nuances as one of the main drives for such a development.

An artificial consciousness should be capable of perceiving what is novel or not, what is original or not, forming a sense of beauty throughout its ontogenetic experience. Aesthetic experience goes hand in hand with emotional experience, surprise, and expectation. We believe that generative models—with all the features that can be built around them, such as predictive processes, prediction error dynamics monitoring, and so the like—can lead to creative abilities in artificial systems and, ultimately, support them in assigning emotional and aesthetic values to activities

and perceptions. An artificial consciousness or a creative predictive machine?

The prehistoric origins of art, according to the archaeologist Steven Mithen (Mithen and Morton, 1996, p. 229), stem from a fluidity of the cognitive domains pertaining to technology, nature, and social life, that allowed our ancestors to leverage symbolic artifacts for cultural development. After many centuries of speculation about sentient machines, we find ourselves in an age in which nature and social life might be fully reflected in our technology, an age in which our technology becomes a social presence. The advantages of this next-step in symbolic culture may lie in the role of consciousness plays in speculation and storytelling, and how these in turn support social cooperation and collaboration (Baumeister and Masicampo, 2010). Consciousness and the assumption of consciousness in each other through theory of mind, is the key to bridging the black boxes of internal cognitive processes we would otherwise be to each other. Human and machine socialization might benefit from similar assumptions.

DHS. Guido?
GS. Yes?
DHS. Are you a zombie?
GS. # @!!

AUTHOR CONTRIBUTIONS

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

FUNDING

GS has received funding from the European Union's Horizon 2020 research and innovation programme under the Marie Sklodowska-Curie grant agreement No. 838861 (Predictive Robots). Predictive Robots is an associated project of the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation) Priority Programme The Active Self.

REFERENCES

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

Baumeister, R. F., and Masicampo, E. (2010). Conscious thought is for facilitating social and cultural interactions: how mental simulations serve the animal-culture interface. *Psychol. Rev.* 117:945. doi: 10.1037/a0019393

Block, N. (2007). Consciousness, accessibility, and the mesh between psychology and neuroscience. *Behav. Brain Sci.* 30, 481–499. doi: 10.1017/S0140525X07002786

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

Boden, M. (2009). "Chapter 13: Creativity: how does it work?," in *The Idea of Creativity* (Leiden: Brill Academic Publishers), 235–250. doi: 10.1163/ej.9789004174443.i-348.74

Bryson, J. J. (2016). "Patiency is not a virtue: AI and the design of ethical systems," in *2016 AAAI Spring Symposium Series* (Cham).

Bryson, J. J., Diamantis, M. E., and Grant, T. D. (2017). Of, for, and by the people: the legal lacuna of synthetic persons. *Artif. Intell. Law* 25, 273–291. doi: 10.1007/s10506-017-9214-9

Carruthers, P. (2016). "Higher-order theories of consciousness," in *The Stanford Encyclopedia of Philosophy*, ed E. N. Zalta (Metaphysics Research Lab; Stanford University, CA) 7.

Chella, A., Cangelosi, A., Metta, G., and Bringsjord, S. (2019). Editorial: Consciousness in humanoid robots. *Front. Robot. AI* 6:17. doi: 10.3389/frobt.2019.00017

Ciaunica, A., and Crucianelli, L. (2019). Minimal self-awareness: from within a developmental perspective. *J. Conscious. Stud.* 26, 207–226.

Ciaunica, A., and Fotopoulou, A. (2017). "The touched self: psychological and philosophical perspectives on proximal intersubjectivity and the self," in *Embodiment, Enaction, and Culture Investigating the Constitution of the Shared World* (Boston, MA: MIT Press Cambridge), 173–192. doi: 10.7551/mitpress/9780262035552.003.0009

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

Damasio, A. R. (2012). *Self Comes to Mind: Constructing the Conscious Brain*. New York, NY: Vintage.

Dennett, D. (1991). *Consciousness Explained (P. Weiner, Illustrator)*. New York, NY: Little, Brown and Co.

Dennett, D. C. (2016). Illusionism as the obvious default theory of consciousness. *J. Conscious. Stud.* 23, 65–72.

Dennett, D. C. (2019). *Will AI Achieve Consciousness? Wrong Question*. Boone, IA: Backchannel, Wired.com.

Du Sautoy, M. (2020). *The Creativity Code: Art and Innovation in the Age of AI*. Harvard, MA: Harvard University Press. doi: 10.4159/9780674240407

Ehrsson, H. H., Spence, C., and Passingham, R. E. (2004). That's my hand! Activity in premotor cortex reflects feeling of ownership of a limb. *Science* 305, 875–877. doi: 10.1126/science.1097011

Eliasmith, C. (2013). *How to Build a Brain: A Neural Architecture for Biological Cognition*. Oxford: Oxford University Press. doi: 10.1093/acprof:oso/9780199794546.001.0001

Escobar-Juárez, E., Schillaci, G., Hermosillo-Valadez, J., and Lara-Guzmán, B. (2016). a self-organized internal models architecture for coding sensory-motor schemes. *Front. Robot. AI* 3:22. doi: 10.3389/frobt.2016.00022

Feldman, H., and Friston, K. (2010). Attention, uncertainty, and free-energy. *Front. Hum. Neurosci.* 4:215. doi: 10.3389/fnhum.2010.00215

Filippetti, M. L., Lloyd-Fox, S., Longo, M. R., Farroni, T., and Johnson, M. H. (2014). Neural mechanisms of body awareness in infants. *Cereb. Cortex* 25, 3779–3787. doi: 10.1093/cercor/bhu261

Filippetti, M. L., and Tsakiris, M. (2018). Just before i recognize myself: the role of featural and multisensory cues leading up to explicit mirror self-recognition. *Infancy* 23, 577–590. doi: 10.1111/infa.12236

Fotopoulou, A., and Tsakiris, M. (2017). Mentalizing homeostasis: The social origins of interoceptive inference. *Neuropsychoanalysis* 19, 3–28. doi: 10.1080/15294145.2017.1294031

Frankish, K. (2016). Illusionism as a theory of consciousness. *J. Conscious. Stud.* 23, 11–39.

Fraser, K. C., Zeller, F., Smith, D. H., Mohammad, S., and Rudzicz, F. (2019). "How do we feel when a robot dies? emotions expressed on twitter before and after hitchBOT's destruction," in *Proceedings of the Tenth Workshop on Computational Approaches to Subjectivity, Sentiment and Social Media Analysis* (Minneapolis, MN: Association for Computational Linguistics), 62–71. doi: 10.18653/v1/W19-1308

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. doi: 10.1038/nrn2787

Gallagher, S. (2000). Philosophical conceptions of the self: implications for cognitive science. *Trends Cogn. Sci.* 4, 14–21. doi: 10.1016/S1364-6613(99)01417-5

Gallagher, S. (2006). *How the Body Shapes the Mind*. Oxford: Clarendon Press. doi: 10.1093/0199271941.001.0001

Gallese, V. (2017). "The empathic body in experimental aesthetics-embodied simulation and art," in *Empathy* (London: Palgrave Macmillan), 181–199. doi: 10.1057/978-1-137-51299-4_7

Garcia-Larrea, L., and Bastuji, H. (2018). Pain and consciousness. *Prog. Neuro-Psychopharmacol. Biol. Psychiatry* 87, 193–199. doi: 10.1016/j.pnpbp.2017.10.007

Gardner, A. (2014). *Ex-Machina*. [*Film*], Universal City, CA: Universal Pictures.

Georgie, Y. K., Schillaci, G., and Hafner, V. V. (2019). "An interdisciplinary overview of developmental indices and behavioral measures of the minimal self," in *2019 Joint IEEE 9th International Conference on Development and Learning and Epigenetic Robotics (ICDL-EpiRob)* (Oslo), 129–136. doi: 10.1109/DEVLRN.2019.8850703

Goff, P., Seager, W., and Allen-Hermanson, S. (2001). *Panpsychism*, Stanford University, CA.

Graziano, M., and Webb, T. W. (2018). "Understanding consciousness by building it," in *The Bloomsbury Companion to the Philosophy of Consciousness* (London, UK: Bloomsbury Publishing), 187. doi: 10.5040/9781474229043.0020

Graziano, M. S. (2016). Consciousness engineered. *J. Conscious. Stud*. 23, 98–115.

Graziano, M. S. (2018). "The attention schema theory of consciousness," in *The Routledge Handbook of Consciousness* (Abingdon, VA: Taylor and Francis), 174–187. doi: 10.4324/9781315676982-14

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

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

Hameroff, S. R., and Penrose, R. (1996). Conscious events as orchestrated space-time selections. *J. Conscious. Stud*. 3, 36–53.

Hayes, S. C. (2002). Buddhism and acceptance and commitment therapy. *Cogn. Behav. Pract*. 9, 58–66. doi: 10.1016/S1077-7229(02)80041-4

Higgins, J. (2020). The "we" in "me": an account of minimal relational selfhood. *Topoi* 39, 535–546. doi: 10.1007/s11245-018-9564-2

Holland, O., and Goodman, R. (2003). Robots with internal models a route to machine consciousness? *J. Conscious. Stud*. 10, 77–109.

Hommel, B., Chapman, C. S., Cisek, P., Neyedli, H. F., Song, J.-H., and Welsh, T. N. (2019). No one knows what attention is. *Attent. Percept. Psychophys*. 81, 2288–2303. doi: 10.3758/s13414-019-01846-w

Horstmann, A. C., Bock, N., Linhuber, E., Szczuka, J. M., Straßmann, C., and Krämer, N. C. (2018). Do a robot's social skills and its objection discourage interactants from switching the robot off? *PLoS ONE* 13, 1–25. doi: 10.1371/journal.pone.0201581

Kac, E. (1997). Foundation and development of robotic art. *Art J*. 56, 60–67. doi: 10.1080/00043249.1997.10791834

Kiverstein, J., Miller, M., and Rietveld, E. (2019). The feeling of grip: novelty, error dynamics, and the predictive brain. *Synthese* 196, 2847–2869. doi: 10.1007/s11229-017-1583-9

Lang, C., Schillaci, G., and Hafner, V. V. (2018). "A deep convolutional neural network model for sense of agency and object permanence in robots," in *2018 Joint IEEE 8th International Conference on Development and Learning and Epigenetic Robotics (ICDL-EpiRob)* (Tokyo), 257–262. doi: 10.1109/DEVLRN.2018.8761015

Lyyra, P. (2010). *Higher-Order Theories of Consciousness: An Appraisal and Application*. Jyväskylä: University of Jyväskylä.

Metzinger, T. (2009). *The Ego Tunnel: The Science of the Mind and the Myth of the Self*. New York, NY: Basic Books (AZ).

Metzinger, T. (2018). "Towards a global artificial intelligence charter," in *Should We Fear Artificial Intelligence* (Bruxelles: EU Parliament) 27–33.

Metzinger, T. (2020). Minimal phenomenal experience. *Philos. Mind Sci*. 1, 1–44. doi: 10.33735/phimisci.2020.I.46

Mithen, S., and Morton, J. (1996). *The Prehistory of the Mind*, London: Thames and Hudson London.

Morse, A. F., De Greeff, J., Belpeame, T., and Cangelosi, A. (2010). Epigenetic robotics architecture (Era). *IEEE Trans. Auton. Ment. Dev*. 2, 325–339. doi: 10.1109/TAMD.2010.2087020

Oudeyer, P.-Y., Kaplan, F., and Hafner, V. V. (2007). Intrinsic motivation systems for autonomous mental development. *IEEE Trans. Evol. Comput.* 11, 265–286. doi: 10.1109/TEVC.2006.890271

Parr, T., Corcoran, A. W., Friston, K. J., and Hohwy, J. (2019). Perceptual awareness and active inference. *Neurosci. Conscious*. 2019:niz012. doi: 10.1093/nc/niz012

Pearl, J., and Mackenzie, D. (2018). *The Book of Why: The New Science of Cause and Effect*. New York, NY: Basic Books.

Pezzulo, G. (2017). *Tracing the Roots of Cognition in Predictive Processing*. Mainz: Johannes Gutenberg-Universität Mainz.

Rochat, P. (1998). Self-perception and action in infancy. *Exp. Brain Res*. 123, 102–109. doi: 10.1007/s002210050550

Rochat, P., and Striano, T. (2000). Perceived self in infancy. *Infant Behav. Dev*. 23, 513–530. doi: 10.1016/S0163-6383(01)00055-8

Schillaci, G., Ciria, A., and Lara, B. (2020a). "Tracking emotions: intrinsic motivation grounded on multi-level prediction error dynamics," in *IEEE International Conference on Development and Learning and Epigenetic Robotics, IEEE ICDL-Epirob* (Valparaiso, IN). doi: 10.1109/ICDL-EpiRob48136.2020.9278106

Schillaci, G., Ritter, C.-N., Hafner, V. V., and Lara, B. (2016). "Body representations for robot ego-noise modelling and prediction. towards the development of a sense of agency in artificial agents," in *Proceedings of the Artificial Life Conference 2016 13* (Boston, MA: MIT Press), 390–397. doi: 10.7551/978-0-262-33936-0-ch065

Schillaci, G., Villalpando, A. P., Hafner, V. V., Hanappe, P., Colliaux, D., and Wintz, T. (2020b). Intrinsic motivation and episodic memories for robot exploration of high-dimensional sensory spaces. *SAGE J. Adapt. Behav*. doi: 10.1177/1059712320922916

Schmidhuber, J. (2006). Developmental robotics, optimal artificial curiosity, creativity, music, and the fine arts. *Connect. Sci*. 18, 173–187. doi: 10.1080/09540090600768658

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

Shergill, S. S., Bays, P. M., Frith, C. D., and Wolpert, D. M. (2003). Two eyes for an eye: the neuroscience of force escalation. *Science* 301:187. doi: 10.1126/science.1085327

Smith, D. H., and Zeller, F. (2017). The death and lives of hitchbot: the design and implementation of a hitchhiking robot. *Leonardo* 50, 77–78. doi: 10.1162/LEON_a_01354

Sneddon, L. U., Elwood, R. W., Adamo, S. A., and Leach, M. C. (2014). Defining and assessing animal pain. *Anim. Behav*. 97, 201–212. doi: 10.1016/j.anbehav.2014.09.007

Spatola, N., and Urbanska, K. (2018). Conscious machines: robot rights. *Science* 359:400. doi: 10.1126/science.aar5059

Thagard, P. (2019). *Brain-Mind: From Neurons to Consciousness and Creativity (Treatise on Mind and Society)*. Amsterdam: Oxford University Press. doi: 10.1093/oso/9780190678715.001.0001

Thagard, P., and Stewart, T. C. (2014). Two theories of consciousness: semantic pointer competition vs. information integration. *Conscious. Cogn*. 30, 73–90. doi: 10.1016/j.concog.2014.07.001

Tononi, G. (2012). *Phi: A Voyage From the Brain to the Soul*. New York, NY: Pantheon.

Wallach, W., Allen, C., and Franklin, S. (2011). Consciousness and ethics: artificially conscious moral agents. *Int. J. Mach. Conscious*. 3, 177–192. doi: 10.1142/S1793843011000674

White, N. (1987). *The Helpless Robot*. Toronto, ON.

Yates, J., Immergut, M., and Graves, J. (2017). *The Mind Illuminated: A Complete Meditation Guide Integrating Buddhist Wisdom and Brain Science for Greater Mindfulness*. New York, NY: Simon and Schuster.

Zahavi, D., and Parnas, J. (1998). Phenomenal consciousness and self-awareness: a phenomenological critique of representational theory. *J. Conscious. Stud*. 5, 687–705.

Zahavi, D., and Parnas, J. (2003). Conceptual problems in infantile autism research: why cognitive science needs phenomenology. *J. Conscious. Stud*. 10, 53–71.

Zangwill, N. (1999). Feasible aesthetic formalism. *Noûs* 33, 610–629. doi: 10.1111/0029-4624.00196

Zoia, S., Blason, L., D'Ottavio, G., Bulgheroni, M., Pezzetta, E., Scabar, A., et al. (2007). Evidence of early development of action planning in the human foetus:

a kinematic study. *Exp. Brain Res.* 176, 217–226. doi: 10.1007/s00221-006-0607-3

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

CONCEPTUAL ANALYSIS
published: 31 May 2021
doi: 10.3389/fpsyg.2021.613587

Whatever Next and Close to My Self—The Transparent Senses and the "Second Skin": Implications for the Case of Depersonalization

Anna Ciaunica[1,2]*, *Andreas Roepstorff*[3], *Aikaterini Katerina Fotopoulou*[4] and *Bruna Petreca*[5]

[1] Institute of Philosophy, University of Porto, Porto, Portugal, [2] Faculty of Brain Sciences, Institute of Cognitive Neuroscience, University College London, London, United Kingdom, [3] Interacting Minds Centre, Aarhus University, Aarhus, Denmark, [4] Research Department of Clinical, Educational and Health Psychology, University College London, London, United Kingdom, [5] Royal College of Art, London, United Kingdom

In his paper "Whatever next? Predictive brains, situated agents, and the future of cognitive science," Andy Clark seminally proposed that the brain's job is to predict whatever information is coming "next" on the basis of prior inputs and experiences. Perception fundamentally subserves survival and self-preservation in biological agents, such as humans. Survival however crucially depends on rapid and accurate information processing of what is happening in the here and now. Hence, the term "next" in Clark's seminal formulation must include not only the temporal dimension (i.e., what is perceived *now*) but also the spatial dimension (i.e., what is perceived *here* or next-to-my-body). In this paper, we propose to focus on perceptual experiences that happen "next," i.e., close-to-my-body. This is because perceptual processing of proximal sensory inputs has a key impact on the organism's survival. Specifically, we focus on tactile experiences mediated by the skin and what we will call the "extended skin" or "second skin," that is, immediate objects/materials that envelop closely to our skin, namely, clothes. We propose that the skin and tactile experiences are not a mere border separating the self and world. Rather, they simultaneously and inherently distinguish *and* connect the bodily self to its environment. Hence, these proximal and pervasive tactile experiences can be viewed as a "transparent bridge" intrinsically relating and facilitating exchanges between the self and the physical and social world. We conclude with potential implications of this observation for the case of Depersonalization Disorder, a condition that makes people feel estranged and detached from their self, body, and the world.

Keywords: self-awareness, touch, altered states of consciousness, depersonalization, body schema, body image, predictive processing

OPEN ACCESS

Edited by:
Francesca Ferri,
University of Studies G. d'Annunzio Chieti and Pescara, Italy

Reviewed by:
Marion Danièle Luyat,
Université de Lille, France
Julian Kiverstein,
Academic Medical Center, Netherlands

***Correspondence:**
Anna Ciaunica
a.ciaunica@ucl.ac.uk

Specialty section:
This article was submitted to Theoretical and Philosophical Psychology, a section of the journal Frontiers in Psychology

Received: *02 October 2020*
Accepted: *19 March 2021*
Published: *31 May 2021*

Citation:
Ciaunica A, Roepstorff A, Fotopoulou AK and Petreca B (2021) Whatever Next and Close to My Self—The Transparent Senses and the "Second Skin": Implications for the Case of Depersonalization. Front. Psychol. 12:613587. doi: 10.3389/fpsyg.2021.613587

INTRODUCTION

In daily life, we experience ourselves as constantly immersed in an ongoing flow of sensory signals (smells, sounds, images, etc.) emerging from both inside and outside our bodies. These sensations and perceptions scaffold both (a) a *sense of self*, i.e., the subjective first-personal "I" or "self," bound to my body and distinct from the world and others (Gallagher, 2000; see Qin et al., 2020 for a recent

review). And, (b) a *sense of presence*, i.e., the feeling that I am immersed and in direct touch with a real world *here and now* (Seth et al., 2011). Perceptual experiences are traditionally regarded as the fundamental point of contact of an experiencing subject with external, mind-independent[1] world situated "out there."

The past decades have witnessed an "embodied turn" in examining perception from a situated, embodied and dynamic viewpoint. The body is not viewed anymore as a mere "material" or "physical" support system for transporting and fueling the mind and brain. Rather minds and brains are designed to support the maintenance and survival of a body within a wider and potentially threatening physical and social environment (Varela et al., 1991; Gallagher, 2005; De Jaegher and Di Paolo, 2007; Thompson, 2007).

In line with this approach, recent work within the influential Predictive Processing (PP) framework suggested that biological agents, such as humans, are "pro-active survival-enabled prediction machines" (Clark, 2013, p. 1) that must optimally maintain their bodily states within the required limits for survival and reproduction purposes[2] (Friston, 2008; Hohwy, 2013; Fotopoulou, 2015; Seth and Tsakiris, 2018). To complete this difficult task, the human brain generates its own internal self- and world-models by extracting statistical patterns of relevant information (Conant and Ross Ashby, 1970; Friston, 2005; Friston and Stephan, 2007). These self- and world-models are constructed moment-by-moment at various levels of the hierarchical processing (Clark, 2013; Hohwy, 2013).

In this view, the main function of perceptual experiences is ultimately geared toward self-preservation. Perceptions are underpinned at the neural level by dynamically shifting generative models "predicting" what is causing incoming sensory events, based on Bayesian probabilistic "guesses" about the likely causes of any incoming sensory information (Friston, 2008; Hohwy, 2013; Fotopoulou, 2015). Incoming sensory inputs are then contrasted or "matched" against learned (or innate) patterns constituting what are called "predictions." When a prediction does not match ongoing sensory input, then a "prediction error" results, which may have the effect of updating the prediction.

It has been proposed that a system's ability to attenuate or to "forget" boring (i.e., predictable, unsurprising) information enables the agent to select "newsworthy" information. For example, when picking a ripe cherry from a tree (Limanowski and Friston, 2020), we seem to be very sensitive to the feel of the cherry, as we touch and grasp it. Yet, we are almost insensitive to the feelings of our arm and eye movements while reaching the cherry. However, these inputs are key in ensuring we successfully pick the ripe cherry, and not the green one next to it.

Importantly, a self-organizing system, such as the human body, is most intimately acquainted with *self*-related signals. This means that the problem the brain has to solve is often "not which sensory evidence to *emphasise*, but which to *attenuate*" (Parr et al., 2018; Limanowski and Friston, 2020, p. 8, original italics). One key observation here is that perceptions and beliefs about the self are unique, in the sense that they are "necessarily transparent" (Limanowski and Friston, 2018, p. 5; Ciaunica et al., 2020). Indeed, it has been suggested that in daily life and under normal circumstances, there is a basic, embodied and pre-reflective sense of self that is "transparent."

As we will see shortly in more detail, the property of "transparency" has been spelled out in a variety of different ways by different theorists (Moore, 1903; Harman, 1990; Tye, 1999; Metzinger, 2003; Sass and Parnas, 2003; Fuchs, 2005; Ciaunica et al., 2020). The basic idea is that transparent processing gives us the feeling of having direct access to what our experiences are about, or directed at. Crucially, this model applies not only to our perception of the external world but also to self-models: "just as a transparent world-model grants the experience of being in immediate touch with the world, a transparent phenomenal self-model... affords the experience of being in immediate relation to a self" (Metzinger, 2003; Limanowski and Friston, 2018, p. 2).

A detailed overview of the PP literature, its advocates and opponents, lies beyond our scope here. In this paper, we retain and build upon the key and largely non-controversial idea that perceptual processing in the here and now is fundamentally geared toward self-preservation (Clark, 2013).

If this is so, then one may pay careful attention to perceptual processing taking place not only in the next second, but literally next-to-my-body. In this paper, we propose to shift the focus on the perceptual processing of proximal sensory inputs (literally "next-to-the-self"), and more specifically on touch which has a key impact on the organism's survival. In line with seminal paper "Whatever next? Predictive brains, situated agents, and the future of cognitive science" by Clark (2013), we suggest that the term "next" should be understood not only at the temporal scale (i.e., what is perceived in the upcoming second) but also from a spatial dimension (i.e., what is perceived literally next to or close-to-my-body).

There are at least two main reasons this shift in focus toward what is happening in proximity to the body. (1) First, as de Vignemont (2018) rightfully notices, the perception of a snake next-to-my foot is essentially different from the perception of the moon in the sky in terms of survival outcome. Indeed, standard accounts tacitly privileged a visuospatial model of perception, which is easily understandable insofar as healthy adults seem to be presented with a continuous visual field devoid of any phenomenological boundary between what is close and what is far (de Vignemont, 2018). Yet, there is growing interest in highlighting the inherent multisensory and proximal nature of perceptual experiences as they unfold throughout the lifespan (Noë, 2004; Ciaunica and Fotopoulou, 2017; Faivre et al., 2017; Fotopoulou and Tsakiris, 2017; Seth and Tsakiris, 2018; Ciaunica and Crucianelli, 2019; Barwitch, 2020).

(2) Second, if our current perceptual experiences are indeed infused and structured by previous "priors" or

[1]For example, Crane and French (2015) note that: "perceptual experience, in its character, involves the presentation (as) of ordinary mind-independent objects to a subject, and such objects are experienced as present or there such that the character of experience is immediately responsive to the character of its objects".

[2]We recently had the chance to see this hypothesis at work during the ongoing COVID-19 pandemic: many of us experienced major difficulties to focus on abstract tasks, our attention being constantly attracted to the concrete conditions that threaten our lives and those of our loved ones.

"expectations," then it becomes crucial to understand how perceptual experiences unfold, dynamically, from a bottom-up developmental perspective. Indeed, in real life, human beings are not emerging as fully-fledged adults–like Athena famously emerged from Zeus' head–but they gradually develop from cells to a human body within another human body (Fotopoulou and Tsakiris, 2017; Ciaunica, 2019; Ciaunica and Crucianelli, 2019; Quintero and de Jaegher, 2020; Ciaunica et al., 2021). As we will see shortly, endorsing a dynamic bottom-up viewpoint points us to the idea that touch, before vision, plays a fundamental role in constituting self- and world-models in early life and beyond (Fotopoulou and Tsakiris, 2017; Ciaunica and Crucianelli, 2019).

Indeed, humans have proximal and "direct sensible acquaintance" (James, 1890) with others' bodies (via skin-to-skin interactions) well before they meet others' minds (via face-to-face and eye gaze interactions) (Ciaunica and Fotopoulou, 2017; Ciaunica et al., 2021). We argue that paradoxically, precisely because tactile proximal inputs are ubiquitous, they tend to be "transparently" processed "in the background," typically unnoticed and taken for granted. Yet, they constitute the fundamental "invisible" roots of our sense of self and sense of presence in the world.

We unpack these ideas below, and we proceed as follows:

In the Perceptual Awareness and the Transparent Experiential Background section, we define the notion of "transparency," and we explore the phenomenological idea that proximal sensory inputs (e.g., olfactory, proprioceptive, tactile, interoceptive) are blended to form a tacit "background existential orientation" (Ratcliffe, 2008). Specifically, we focus on tactile experiences as pervasive yet essential components of our everyday life, forming a baseline and transparent experiential background.

The Getting Perception off the Ground: The Basic Proximal Senses section motivates the claim that tactile experiences may precede more sophisticated forms of detached and distal perceptual awareness—such as visuospatial perception (e.g., seeing an apple). Indeed, well before humans are able to recognize themselves in a mirror and to perceive themselves from a visuospatial distal perspective, they experience themselves and their surroundings via proximal senses. We review empirical developmental findings pointing to the key role of touch in constituting perceptual experiences in early life and beyond.

In the (Un)Covering our Body—Extending the Transparency to the "Second Skin" section, we tackle a relatively overlooked aspect of our tactile experiences, namely, the fact that with rare exceptions, humans in modern societies spend most of their lives having their bodies closely enveloped by clothes. We focus on tactile experiences that extend to the immediate objects/materials that envelop closely our skin. Following the seminal "extended mind" thesis (Clark and Chalmers, 1998; Kirchhoff and Kiverstein, 2019), we argue that these materials may be conceived as a "second skin" or "extended skin" that underwrite what we will call here "extended body-image" and "body-schema" (cf. Gallagher, 2005).

Finally, in The Transparent Bridge—Bodily Connectedness Through Proximal Senses section, we suggest that tactile experiences–mediated by the skin and the "extended skin" – may be viewed as a "transparent bridge" intrinsically relating and facilitating exchanges between the self and the physical and social world. We build upon the observation that the skin and tactile experiences have an inherent dual function: it simultaneously separates and relates our bodily self to the physical and social world. Touch mediates indeed our self-presence in the world. We thus hypothesize that close tactile engagements with the social and physical environment play a fundamental role in shaping both our sense of self and sense of presence or immersiveness in the world. We briefly discuss some potential implications of this observation for the case of Depersonalization Disorder, a condition that makes people feel estranged and detached from their self, body, and the world. We conclude with some potential implications of our hypotheses for designing bodily and interactive interventions aiming to alleviate the symptoms of depersonalization. If our hypotheses are correct, then it is paradoxically by inviting depersonalization people to "forget" their self and to get closer to the world and others that they may start to get closer to their "lost" self.

PERCEPTUAL AWARENESS AND THE TRANSPARENT EXPERIENTIAL BACKGROUND

Perceptual awareness is typically described as having a polarized subject–object structure whereby an experiencing subject perceives an "external" world/object "out there." Traditionally, it has been proposed that one becomes aware of one's self when one is able to perceive oneself as an object of one's awareness and to see oneself through others' eyes (Carruthers, 1996). For example, when I see myself in a mirror, I recognize myself as an individual distinct from the world and others. But also, I can see myself as others see me, from a third-person perspective.

However, as James (1890) pointed out, before we see ourselves as others see us, we perceive ourselves and our bodies through feelings. The most fundamental way to get to know oneself is through feelings, not thought: "For this central part of the Self is *felt*... It is at any rate no *mere ens rationis*, cognized only in an intellectual way, and no mere summation of memories or mere sound of a word in our ears. It is something with which we also have *direct sensible acquaintance*... when it is found, it is *felt*; just as the body is felt" (James, 1890, p. 298–299 original italics; bold our emphasis).

Indeed, a long-standing phenomenological tradition pointed out that among the objects that an experiencing subject perceives in the world throughout her life, the body has a special status (Merleau-Ponty, 1962). For example, Husserl famously wrote that the body is not just an object that is perceived but also that through which we perceive: "The Body [*Leib*] is, in the first place, the *medium of all perception*; it is the *organ of perception* and is *necessarily* involved in all perception" (Husserl, 1989, p. 61). The body follows the experiential subject everywhere, like a shadow, for better and for worse (Legrand, 2006).

In line with these ideas, it has been proposed that all experiences include a "background existential orientation" constituted by bodily feelings: "the feeling *is* the way in which one finds oneself in the world and the way in which one finds oneself

in the world participates in all experience, albeit as something that is usually pre-reflectively taken for granted" (Ratcliffe, 2008, p. 140).

The idea of an existential bodily background has been highlighted in neuroscience research. For example, Damasio writes:

> "I am postulating another variety of feeling which I suspect preceded the others in evolution. I call it *background feeling* because it originates in 'background' body states rather than in emotional states. It is not the Verdi of grand emotion, nor the Stravinsky of intellectualized emotion but rather a minimalist in tone and beat, the feeling of life itself, the sense of being. (...) The background feeling is our image of the body landscape when it is not shaken by emotion. (...) I submit that without them the very core of your representation of self would be broken" (Damasio, 1994, p. 150–151, our italics).

The lived body is the feeling body that remains tacitly, and one could say "transparently" in the background[3]. In the remainder of this paper, we build upon the idea that these bodily feelings constitute a pervasive and "transparent" experiential background, which can be attended or altered, i.e., "opaque" or "broken" (Fuchs, 2005; Ciaunica et al., 2020).

Transparency can be intuitively grasped via the "window" metaphor: a clear and transparent window glass or sliding door can give us the illusion of an unmediated access to the outside world. However, although we subjectively feel that we are directly in touch with one's inner self and the outer world, in reality, our experiences are *mediated* through certain states or processes that are transparent, pervasive, and tacitly taken for granted (Fuchs, 2005; Ciaunica et al., 2020).

As we saw earlier in the Introduction section, the notion of "transparency" has been theoretically spelled out in different ways by different theorists (Moore, 1903; Harman, 1990; Tye, 1999; Metzinger, 2003), and a detailed review of these accounts lies beyond the scope of this paper. Here, we restrict our focus on the property of transparency of perceptual experiences. To get a clear intuitive grip on this idea, consider the following example: I am on a street, and I hesitate to turn my head left or right. Eventually, I turn my head left, and I realize that I chose this direction, because I have perceived with the corner of my eye the label of a French bistro. Suddenly, I realize that I am hungry, and the reason why I turned my head into that direction is because my brain, concerned by my bodily survival, detected some relevant "transparent" background information, that is now brought forward, at the surface of my awareness (the smell of delicious French cuisine, say). Consequently, I engage in self-regulatory behavior, and I walk toward the bistro and order escargots.

Here, we restrict our focus on the property of transparency of experiences (and leave aside the property of transparency of mental representations[4] or processes). Note that it is important to distinguish, in our view, between (a) what is non-conscious but can be in principle consciously attended (brought into the focus of the attention; e.g., the visual stimuli of a bistro label or the interoceptive signals of my hungriness); and (b) what is non-conscious and cannot be consciously attended (e.g., the firing of my neurons while I perceive the French bistro; or the metabolic exchanges between my blood and my organs). Only the former information, but not the latter, can be processed "transparently" because only (a) but not (b) can be attended, and thereby become "opaque."

In what follows, we explore the idea that our tacit "background existential orientation" is essentially underpinned by proximal sensory inputs (e.g., olfactory, proprioceptive, tactile, and interoceptive) that are blended to form an experiential "transparent" background. These signals are so pervasive and yet so essential to our being in the world that they seem to form a baseline and taken for granted experience. Specifically, we focus on touch[5] and tactile experiences for three interrelated main reasons. First, touch is mediated by the skin, the oldest and widest organ in terms of dimensions and functions (Montagu, 1971; Field, 2001; see Gallace and Spence, 2010 for a review). By providing the organism with the most primitive means to "meet" and perceive the world, tactile experiences may constitute thereby the most ubiquitous and basic experiential background.

Second, the skin mediates the boundary between the self and the outer world, and as such, tactile experiences display an inescapably dual "touchant/touché" structure (Merleau-Ponty, 1962). By gaining information about the world via touch, the subject inherently gains information about her "self" too. Indeed, while proprioception, kinaesthesia, and interoceptive of visceral inputs are phenomenologically inextricable components of the existential background feeling, tactile experiences because they play a special relational or dual role: "tactual perception" (Ratcliffe, 2013) is closely associated with or partly constituted by perception of one's body: one cannot perceive the world tactually without perceiving oneself in the process. As Merleau-Ponty (1962, p. 316) famously pointed out, while vision "presents us with a spectacle spread out before us at a distance," in perceiving the world through touch, "I cannot forget in this case that it is through my body that I go to the world" (1962, p. 316). By directly mediating the boundary between body and world, the skin inescapably distinguishes yet relates body and world, as the two faces of the same coin. As Martin notes (Martin, 1992, 1993, 1995), tactile feelings can simultaneously be perceptions of body and/or of world.

[3]As Merleau-Ponty writes: "in so far as it sees or touches the world, my body can be neither seen nor touched. What prevents its ever being an object, ever being 'completely constituted' is that it is that by which there are objects. It is neither tangible nor visible in so far as it is that which sees and touches." (Merleau-Ponty, 1962, p. 92).

[4]For example, some theorists argue that while we typically have access only to the mental representation's intentional content (something in the world which it is about) without noticing its non-intentional carrier properties (Moore, 1903; Harman, 1990; Tye, 1999), the process itself of constructing inner representations can become available to our introspective attention. Whenever we consciously direct our attention introspectively inwards, so to speak, the transparent processing of mental representations (typically taken for granted and hence "invisible") becomes "opaque," that is, "visible" and available to our attention (cf. the window metaphor described above).

[5]We do not claim that touch is the most fundamental proximal sensory channel, rather we outline its special role as a relational sense par excellence. We are grateful to one anonymous reviewer for pressing clarification on this point.

Lastly, but importantly, touch plays a key exploratory and social bonding role, which confers on it a sense of "closeness": we touch things to make sure they are real, and we touch people to make direct and close contact with them. As Fulkerson (2014) notes, what distinguishes tactual experience from the other senses is the fact that the former involves "exploratory binding"; it relies upon actively manipulating (and being manipulated by) the environment.

Up to now, we have motivated the claim that the body–world relationship mediated via tactile perception remains tacitly and ubiquitously in the background and the sense of self may be inseparable from this relationship. In the next section, we suggest that close tactile experiences may precede developmentally more sophisticated forms of detached and distal perceptual awareness—such as visuospatial perception (e.g., seeing an apple). Well before humans are able to recognize themselves in a mirror and to perceive themselves from a visuospatial distal perspective, they experience themselves via proximal senses, such as tactile experiences.

GETTING PERCEPTION OFF THE GROUND: THE BASIC PROXIMAL SENSES

There is a growing consensus in philosophy, psychology, and cognitive neuroscience that multisensory information about the body plays a central role in structuring our basic sense of self (Gallagher, 2000; Blanke and Metzinger, 2009). Specifically, the interplay and coupling between (a) exteroceptive (e.g., vision, audition) and (b) interoceptive senses (e.g., temperature, pain, cardiac signals, breath, etc.) is a key component of our sense of self (Park and Blanke, 2019). More recently, the PP framework proponents argued that the basic experience of being a self is the result of an ongoing inferential process based on a generative model centered onto the bodily self (Apps and Tsakiris, 2013; Hohwy, 2013, 2020; Limanowski and Blankenburg, 2013; Seth, 2013; Limanowski and Friston, 2018, 2020).

Biological agents, such as humans, act as "self-evidencing" systems aiming at maximizing evidence for their self-model as they minimize prediction errors (Hohwy, 2014). In short, self-organizing systems need to constantly "check" and "prove" themselves that they exist in a dynamic, constantly changing and noisy world. The self is an inferred model of endogenous, deeply hidden causes of behavior. It has been proposed that self-modeling can be described as a two-step process: (i) first finessing a model of the self and then (ii) engaging that model in action, which spirals into further finessing of the model, and further action, and so forth (Hohwy and Michael, 2017).

However, one basic yet overlooked aspect of current embodied and PP approaches in both philosophy and cognitive neuroscience is that brains (and minds), and human bodies, first develop *within* another human body. The most basic models of perceptions and actions emerge already *in utero* (Ciaunica, 2019; Ciaunica and Crucianelli, 2019; Quintero and de Jaegher, 2020; Ciaunica et al., 2021). Crucially, while not all humans will have the experience of being pregnant or carrying a baby, the experience of *being carried and growing within another person's body* is universal[6] (Ciaunica et al., 2021).

In the remainder of this section, we provide evidence illustrating that it is touch (and not vision) that is one of the first of our senses to develop and affords us thereby with our most basic and earliest means of "meeting" and perceiving both the self and the external world (Ciaunica and Fotopoulou, 2017; Fotopoulou and Tsakiris, 2017; Ciaunica and Crucianelli, 2019).

In the womb, fetuses spend a significant amount of time in tactile exploration of the boundary between innervated and non-innervated regions (Mori and Kuniyoshi, 2010; Piontelli, 2014; Hata, 2016). For example, fetuses frequently touch certain body areas, such as the lips, cheeks, ears, and parietal bone, creating a self-stimulatory pattern, which enhances innervation. Importantly, when the fetus touches the forehead, innervation increases, and the boundary migrates (Piontelli, 2014). This allows the fetus to move on in touching a new innervated boundary, and the cycle repeats until the whole body is fully innervated (Piontelli, 2010; Delafield-Butt and Gangopadhyay, 2013). Additionally, when the fetus touches itself, the placenta, or a co-twin, it develops different kinematic and tactile patterns that emerge, which differ in pressure, acceleration, and directedness (Hata, 2016).

For example, Castiello et al. (2010) investigated the kinematic profiles of movements in five pairs of twin fetuses by using four-dimensional ultrasonography during two separate recording sessions carried out at the 14th and 18th weeks of gestation. They showed that by the 14th week of gestation, twin fetuses do display not only movements directed toward the uterine wall and self-directed movements but also movements specifically aimed at the co-twin, the proportion of which increases between the 14th and 18th gestational weeks. They also noted similar kinematic profiles for movements directed toward the co-twin and self-directed movements aimed at the eye-region, i.e., the most delicate region of the body. They concluded that performance of movements toward the co-twin is not accidental: already starting from the 14th week of gestation, twin fetuses execute movements specifically aimed at the co-twin. This has been reported in singleton pregnancy as well, where kinematic studies seem to suggest that motor planning is in place by 22 weeks of gestational age (Zoia et al., 2007).

Moreover, it has been shown that maternal touch of her own abdomen increases arm, head, and mouthing movements in the fetus (Marx and Nagy, 2015), and that maternal touch has more impact than maternal voice in the fetus' movements. Notably, tactile interactions require the "toucher" and "touched" to be physically proximal, to "share" the experience of touch (passive or active), and it is often accompanied by a cascade of other sensorial information, such as the smell of the other person, the sound of the tactile contact on the skin (think of the "noise" made by a kiss), and temperature of the other body. Given the richness of information provided by tactile interactions, it has been hypothesized (Ciaunica and Fotopoulou, 2017; Fotopoulou and Tsakiris, 2017) that social touch might represent a fundamental

[6]Throughout this paper, we will use the terms "pregnant person" and "mother" interchangeably, irrespective of their self-identified gender.

step in the development of both self- and other-awareness, as well as self-other distinction (see McGlone et al., 2014 for a review).

Specifically, one type of pleasant touch, the so-called affective, or sensual touch (i.e., slow, caress-like touch mediated, among other, more classic tactile fibers, by the C Tactile afferent system, Löken et al., 2009; Kirsch et al., 2020), has been re-defined as an interoceptive[7] modality since its primarily functional role seems to be to provide information about the homeostatic and emotional effects of touch, rather than the properties of what or who is being touched. It has also been shown that the activation of these CT receptors on the skin may specifically relate to the positive consequences of interpersonal touch (see McGlone et al., 2007), such as reducing feelings of social exclusion (von Mohr et al., 2017), soothing pain (Krahé et al., 2016; von Mohr et al., 2018), and communicating social support (Kirsch et al., 2018). Indeed, related findings suggest that this system may be specialized for processing not only affective touch but also specifically *social affective touch* from early on in life (see Morrison et al., 2010 for a review). Specifically, the fetus is entirely covered in fine hair (i.e., lanugo hairs), and it has been suggested that fetal movement in the amniotic fluid might directly stimulate CTs afferent, which are known to activate the hypothalamus and insular cortex, promoting an anti-stress effect via realize of oxytocin and stimulating the fetal growth (Bystrova, 2009).

A second key example of early sensorimotor coordination is self-touch and tactile perceptions. While a significant body of research focused on the effect of auditory inputs (such as maternal voice or music) on the fetal development, it is important to bear in mind that in the womb, the most developed sensory systems in fetuses are the tactile and olfactory ones, not the visual or auditory ones. This suggests that the most basic way to perceive oneself and the world is mediated through proximal modalities (Ciaunica and Fotopoulou, 2017), such as touch, interoception, and olfaction (Ciaunica and Crucianelli, 2019).

After birth, infants receive constant and proximal tactile stimulations, which significantly reduce infant's stress and increase positive affect (Stack and Muir, 1992; Bellieni et al., 2007). Affective touch is believed to play an important role in the creation and maintenance of social bonding[8] (see Morrison et al., 2010 for a review) and more recently to the sense of body ownership (Lloyd et al., 2013; van Stralen et al., 2014) and self-identity (Panagiotopoulou et al., 2017).

We can summarize the argumentative backbone developed so far as follows: if we grant the premise that the perception of incoming "whatever next" (Clark, 2013) sensory information emerges from the processing of "whatever before," i.e., priors or "expectations," then one needs to step back and have a closer look at how perception gets off the ground from the outset. If we endorse a bottom-up developmental perspective, then empirical work points us to the idea that (discriminative and affective) tactile inputs play a fundamental role in constituting perceptual experiences in early life and beyond.

In the next section, we turn our attention to the special relationship between touch, body, and what is next-to-our-skin. We introduce the key notion of "the second skin" or "extended skin" and explore them in relation to materials situated close to our skin, i.e., our clothes. We claim that they constitute what one may call, following Gallagher (2005), "extended" body-image and body-schema. We turn to this discussion now.

(UN)COVERING OUR BODY—EXTENDING THE TRANSPARENCY TO THE "SECOND SKIN"

Most of the time, in our daily lives, our body is covered with clothes. Textile are pervasive material objects enveloping closely our skin. Clothes may thus play a peculiar role in our bodily experiences. First, they seem to be at the same time visible and invisible. Visible because they are tangible and on display on our bodies and the bodies of others. Invisible, because we tend to "forget" them in the background and to process them "transparently" as defined above. Second, materials that envelop our body allow individuals to closely relate to and exchange with the external physical and social world.

Throughout human history, individuals ingeniously made use of various materials close to their bodies—for protection, comfort, and aesthetic aspects. Particularly, textiles and clothes, which throughout the "long and intimate" (O'Connor, 2010) connection with people, were key in symbolic representations of self (e.g., identity, gender, age, class, political views, as well as a variety of social values; Guy et al., 2001; Weber and Mitchell, 2004; Mida and Kim, 2015), in the reproduction of social order (e.g., in making social differences visible, see Entwistle, 2000; Breward, 2003; Crane, 2012), in embodying culture (memory, history, and identity), and in "transforming, protecting, and healing the human body" (O'Connor, 2010; see also Hansen, 2004; Brumfield, 2006). Many times, textiles have been used with varying degrees of explicitness as a medium to record history or to "provide a focus for the expression of conflict or reflect commentary on current affairs" (Henderson, 1990). The latter can be noticed, for example, in more codified manners, such as in Ganseys[9], or in a very explicit manner in t-shirts portraying rock music bands or political messages, a style very much influenced by Vivienne Westwood's collaboration with Sex Pistols in the 1970s. Hence, we can view clothing as "the furniture of the mind made visible" (Laver in Barnard, 2007, p. 2).

[7]Interoception (the perception of bodily visceral signals) is taken to be uniquely related to the generation of subjective feelings, informing the organism regarding its levels of arousal and bodily needs (Craig, 2009; Seth, 2013). Even more importantly, and contrary to classic views of interoception as "the perception of the body from within," the current notion of interoception is tightly linked to homeostasis. Within this framework, interoceptive signals are considered crucial in informing the organism regarding the homeostatic state of the body in relationship to experiences originating from both within the organism (e.g., cardiac and respiratory functions, digestion, hunger, and thirst) or outside it (e.g., taste, smell, affective touch, and pain).

[8]Substantial literature in the context of attachment formation also supports the facilitating role of touch in establishing the social bond between infant and caregivers (Ainsworth, 1979; Weiss et al., 2000).

[9]A gansey (or guernsey) is a type of knitted woollen jumper, historically known for being worn by fishermen around the coast of Britain. Their unique patterns represented a village and a family, so that in the case of a shipwreck or accident, the bodies could be identified and returned to their family for care or burial.

This long and intimate relationship also endowed textiles with magical, and at times humanlike, characteristics in people's imaginaries. Textiles may be perceived as being "alive" (e.g., designers describe textiles as an animated, living thing with whom they engage in conversation to define what a textile wants to be as a garment; for a full account, see Petreca et al., 2015), or, as observed by Henderson (1990, p. 1,588) may be seen to "have supernatural properties" (e.g., the invisibility cape in Harry Potter novels), or "may be used as a metaphor for life" (e.g., "Penelope rips out her weaving nightly in order to arrest the flow of time"), or even to be "imbued with the characteristics of their wearers or impart special qualities to them" (e.g., the widespread culture of celebrities' clothes auctioning, where pieces are sold at exorbitant prices[10]).

The clothes lie at the interface of the body and its social presentation, and they "signify to the wider world who and what a person is trying to express" (Salgado-Montejo et al., 2018). This symbolic level is rooted on the physical touch basis of the experience of wearing, and we saw earlier that this long and intimate, proximal experience grants textiles "magical" perceptions of "being alive" and "being imbued with the characteristics of their wearers." Indeed, clothes have a unique capacity to provide immediate feedback to an individual, the wearer, affirming their sense of identity at a directly embodied level (Twigg and Buse, 2013). Clothes can express largely unconscious aspects both of an individual and of a group psyche (Salgado-Montejo et al., 2018) and form a part of our non-verbal communication, or what is defined by Lurie (2000), as a "universal tongue."

What we wear reflects our personality, or undressed self, and therefore, it is a material manifestation of our beliefs and behaviors, or "how we think and how we live" (Salgado-Montejo et al., 2018, p. 2): "enclothed cognition is a bidirectional phenomenon, with the wearer selecting clothes that reflect his/her personality and clothing having an influence on cognitive processing and behavior" (Salgado-Montejo et al., 2018, p. 2). For example, it has been suggested that people are efficient and effective in forming impressions of someone's personality via clothing, specifically shoes (Gillath et al., 2012). Furthermore, clothing can influence self-perception. People wearing formal clothing considered themselves as competent and rational (Hannover and Kühnen, 2002; Peluchette and Karl, 2007) and also have enhanced abstract cognitive processing (i.e., comprehensive mental representations) and increased perceived social difference and power (Slepian et al., 2015).

In what follows, we suggest that the clothes may be described as a "second skin" and as such they are an essential component of the transparent experiential background defined in the Perceptual Awareness and the Transparent Experiential Background section. They "extend" the body's schema and image. In his seminal formulation, Gallagher defined (i) "body image" as "a complex set of intentional states and dispositions, perceptions, beliefs, and attitudes, in which the intentional object is one's own body" (Gallagher, 2005, p. 25). By contrast, (ii) the "body schema" is a "non-conscious performance of the body, i.e., a performance that is not an intentional object present to my consciousness" (Gallagher, 2005, p. 548).

Interestingly, he points out that the body-schema can include tools and artifacts as well:

> "the carpenter's hammer becomes an operative extension of the carpenter's hand, or as noted, the body schema extends to the feather in the woman's hat (see Gorman, 1969, p. 15). The body schema is an active, operative performance of the body, rather than a copy, image, global model, or conception of the existing parts of the body. The schema is the body as it actively integrated its position and responses in the environment. (...) The body schema, understood this way, is not the perception of 'my' body; it is not the image, the representation, or even the marginal consciousness of the body. Rather is it precisely the style that organises the body as it functions in communion with its environment" (Gallagher, 1986, p. 548–549).

The extended mind thesis famously defended by Clark and Chalmers (1998) claims that the cognitive processes can be "offloaded" or "extended" to reach beyond the boundaries of individual, such as to include as proper parts aspects of the individual's physical and sociocultural environment (see Kirchhoff and Kiverstein, 2019 for a recent review). Kirchhoff and Kiverstein (2019) use the example of a spider and its web: the latter becomes transparent for the spider and perceives its prey through the web. But how about bodily sensory processes? Can they be "offloaded" or "extended" to reach beyond the boundaries of the skin?

We suggest that some perceptual bodily experiences can be extended through (i.e., transformed into) components that reach beyond the sensory organ's states (e.g., the skin) into what one may call an "extended skin" or a "second skin." Hence, the transparency of tactile experiences may be "offloaded" into the materials close to the skin. The notion of "offloading" refers to the use of artifacts (here clothes) to manage and convey information about the self. For example, clothes may be regarded as a key component of what one may call an (i) "extended" body-schema (e.g., regulating body temperature via warm clothes) and (ii) "extended" body-image, contributing to the constitution of one's social (narrative) self via social signaling (e.g., wearing a red T-shirt to signal that we support Arsenal football club; see Colombetti, 2016).

The idea is that clothes are typically processed transparently, in the background. This is because, by being so close to our skin, they "inherit," so to speak, the pervasiveness and transparency of tactile experiences, which then "extends" to the clothes. The skin being the largest of our sensory organ and given that most of the time, most of our bodies (and hence skin) are covered by clothes, this means that we may tacitly "sense" the world through this interface. We can thus speak of an "extended transparency."

[10]This culture has reached the level of granting clothes and accessories even with world records, such as the most expensive ever pair of trainers, which was worn and signed by Michael Jordan, and sold at $560,000 at a Sotheby's auction held on May 17, 2020. For a variety of examples, see a recent feature. Available online at: <https://www.businessinsider.com/most-expensive-celebrity-memorabilia-items-sold-at-auction?r=US&IR=T#a-pair-of-nike-air-jordan-1s-game-worn-by-michael-jordan-in-1985-and-signed-by-the-player-sold-for-560000-in-a-sothebys-auction-1>.

There is a robust body of work illustrating that our representation of our own body and our bodily sensations does not end on our skin; instead, the proximal space around the body, known as the peripersonal space (PPS), is included in our multisensory representation of the body (see Serino, 2019 for a recent review). Early neurophysiological studies in primates identified a specific population of multi-sensory neurons that selectively responded to external stimuli, but only when the somatosensory, visual, or auditory stimulation occurred close to and not far from the body (Rizzolatti et al., 1981). Subsequent neuropsychological (e.g., di Pellegrino et al., 1997; Làdavas, 2002; di Pellegrino and Làdavas, 2015), psychophysiological (e.g., Maravita et al., 2002, 2003), and neuroimaging (e.g., Bremmer et al., 2002; Blanke et al., 2015; Cléry et al., 2015) studies showed that the processing of tactile information is strongly affected by external stimuli presented close to the body, further cementing the notion that a similar system in humans exists.

This space representation is thought to be dynamic rather than static, so as to maximize the processing of relevant events for the self, expanding, and shrinking depending on environment specificities. Moreover, sociological and anthropological studies have revealed the role of PPS in the type and strength of social and cultural interactions (Hall et al., 1968; Sommer, 1969; Hall, 1990; Felipe and Sommer, 2017). More recently, social neuroscience brought these ideas together to show that interacting proximally with another person influences how we perceive not only our own body but also the space around it, depending on our precise relationship with that person and how much we trust them (Heed et al., 2010; Teneggi et al., 2013).

A detailed discussion of the fascinating topic of the role of PPS in constituting self-representations and our sense of self lies beyond the topic of this paper[11]. Here, we retain the idea of a literally "close relationship" between clothes and bodily self-experiences.

Note that there is an important distinction between "extended body schema" in tool as opposed to clothes (daily) experiences[12]. Take for example a blindman's stick. A successful tool use in this case implies that the individual successfully *attends to* the tip of the stick in order to explore and perceive its environment: she/he can "see" with the stick. The same goes for the carpenter's hammer: he needs to attend to the hammer's contact with the nail in order to hit it correctly. In some cases, the tools may become transparent too[13]: think of a person expertly typing on a laptop keyboard while focusing entirely on the letters appearing on the screen. However, the link between the keyboard and the letter appearance needs to be available for attention, in order to correctly perform the task. If the letters appear on the screen with a delay after I have hit the keyboard, then the transparency breaks down, and I may need to bring my attention to the keyboard itself. In other words, there is an instrumental, functional role in the tool use that requires attention in some way or another.

By contrast, everyday clothes may blend into the background unnoticed. They also need to be comfortable such as we can afford to dis-attend and "forget" about them. In other words, to process them via what one may call an extended transparent experiential background. Unless these proximal tactile experiences become unpleasant or uncomfortable (rough materials, itchy, dry skin), in which case they become "opaque" and emerge at the surface of our awareness. There are cases where clothes are deliberately brought into the foreground, to signal power and social status: think of a king wearing a heavy ermine mantle. The key idea here is that by mediating the boundaries between our bodies and the environment, clothes also seem to have an ambivalent role: they are simultaneously *covering* the intimacy of the self while *uncovering* the self to signal status and power to the public eye.

Up to now, we focused on proximal perceptual experiences as mediated by touch. We took the example of clothes close to our body and suggested that tactile experiences mediated via the skin and the "second skin" become pervasive, transparent and allows what Gallagher called above a "communion with its environment".

In the last section of our paper, we examine some of the implications of shifting the focus from visuo-spatial to proximal perceptual experiences in relation to our sense of self and sense of presence in the world. Specifically, we briefly describe the case of Depersonalization, a condition that makes people estranged and detached from their self, body, and world. We discuss some empirically testable hypotheses, to be developed in the future work, highlighting the key role of proximal tactile interactions with the environment on one's sense of self and sense of presence in the world. If correct, our hypotheses may fruitfully contribute to designing body-based and dynamic interventions helping to alleviate feelings of self-estrangement and social alienation in people suffering from depersonalization symptoms.

THE TRANSPARENT BRIDGE—BODILY CONNECTEDNESS THROUGH PROXIMAL SENSES

While traditionally theorists have been concerned mainly with the subject–object polarity in describing perceptual experiences, throughout this paper, we aimed at drawing attention at what is happening *in between* these two poles. For this reason, we proposed to shift our focus from visuo-distal perception (easier to structure around a subject/object axis) to proximal perception (and particularly on touch), where the key role of mediating boundaries becomes more evident. Contrary to the standard approach that views the skin (and tactile experiences) as a mere border separating the self and world, here we propose that the skin (and its extended version, "the second skin," i.e., the clothes) simultaneously and inherently distinguishes *and* connects the bodily self to its environment.

We motivate this shift in light of the new and highly influential PP theories of perception. As we saw in the Perceptual Awareness and the Transparent Experiential Background section, one key idea behind embodied cognition and PP accounts is

[11] See de Vignemont and Iannetti (2015) for an interesting discussion on different forms of PPS.

[12] We are grateful to an anonymous reviewer for pressing clarification on this point.

[13] We are grateful to an anonymous reviewer for pressing extra clarification on this point.

that perceptual experiences are deeply depended on dynamic engagements and reciprocal interactions with others and the environment (Gibson, 1979; Schilbach et al., 2013). If this approach is correct, then our sense of self and presence or immersiveness in the world may crucially depend on how we relate with our social and physical surroundings. To put it in a slogan: the more we interact with the world and others, the more we gain information about ourselves.

Crucially, human bodies do not emerge in a vacuum. Indeed, one fundamental yet overlooked aspect of our embodiment is that the human organism emerges and develops *within* another human body. While the experience of pregnancy has been traditionally linked to a certain category of individuals (pregnant persons), to date, all humans shared their bodies with the body of another person. Consequently, one needs to adapt and extend the notion of embodiment to reflect the fundamental and universal body-within-a-body case, what it has been termed co-embodiment (Ciaunica et al., 2021).

The regulation of the two agents' states–and particularly the need to maintain physiological stability despite a fluctuating and unpredictable environment–is actively negotiated between the two organisms that share for a given amount of time (typically 9 months) common bodily and environmental resources. Importantly, humans start to perceive themselves and to relate to their close environment, already in the womb, via proximal senses (such as touch) way before vision. After birth, human babies are unable to self-regulate their own homeostatic balance; hence, they fundamentally depend on close physical interactions with dedicated caregivers for survival (Ciaunica and Fotopoulou, 2017; Fotopoulou and Tsakiris, 2017; Ciaunica et al., 2021). As toddlers and adults, we actively explore our proximal surroundings and touch objects to make sure they are real. Since all humans have started their experiential journey as babies within another's body, the co-embodied and relational aspect of perceptual experiences may constitute the basis of our fundamental sense of self.

While there is growing awareness that perceptual experiences are multisensory in nature (Faivre et al., 2017), in this paper, we narrowed our focus on the inescapably dual *touchant/touché* structure of the tactile experiences. This reciprocity and duality place touch in a key position highlighting the inherent relatedness of our bodies with its close physical and social world. We propose to label this overlooked relational background mediated by the proximal senses–and particularly by touch–as a "transparent experiential bridge" that is there without us being aware that it is there. One key question that further work needs to address is the nature of the link between the phenomenological sense of self and this experiential bridge. How does this relationality of our experiences at their most primitive stages fits within the traditional picture of pre-reflective self-awareness or minimal self as being fundamental to all experiences?

While we cannot do justice to this question in this paper, the remainder of this section, we would like to suggest that in daily life, healthy humans constantly use this transparent bridge to communicate and to relate with the world and others, without paying too much attention to its "invisible" and "transparent" underlying structure and pillars, so to speak. However, the importance of this key relatedness becomes apparent especially when this "transparency" breaks down and people start feeling unreal and disconnected from their selves, bodies, and the world[14].

As Ratcliffe rightly points out, people usually talk about a self as distinct from the body whenever they feel something is disrupted or amiss with the self: something that has gone wrong or something that is lacking. To use our initial metaphor: we notice more easily the existence of a transparent window when its glass cracks (Ciaunica et al., 2020). When the mediating function of transparent senses gets disrupted, people may feel disconnected from their self, body, and world (Ciaunica et al., 2021).

Initially described by Dugas in 1898 (Sierra and Berrios, 1998), Depersonalization/Derealisation Disorder (DP/DR) is a condition characterized by profound alterations of one's sense of self (Sierra and David, 2011), typically inducing distressing feelings of detachment or estrangement from one's self (depersonalization) and/or one's surroundings (derealization) (DSM IV-TR fourth edition, text revision 2000)[15]. DP/DR typically co-occurs in association with highly traumatic events or as symptoms of anxiety, panic, and depression.

These dramatic alterations are typically experienced as a "split" or a "fracture" between an external observing agent and an observed self, body, and world: "My perception felt as though it had been drawn back inside my head, almost as though I was looking at the world from the back of my head, and could see the back of my own eye sockets. (...) Essentially, it felt like there was a divorce or fracture between the world and me so that although my body was still in the world, my mind was only an observer" (Ciaunica et al., 2020, p. 6).

The experienced self-split or self-detachment occurs on multiple levels, as it is associated with (a) detachment from one's body or body parts (low-level sensory and bodily aspects of the self); (b) detachment from one's subjective feelings and emotions (experiential aspects); and (c) disconnection from one's personal stories, memories, thoughts, and future plans, often described by sufferers as a lack of a narrative[16] or a "plot" in one's life (Ciaunica and Charlton, 2018). The overall impact of this "self-split" makes

[14]For example, the current social isolation provoked by the COVID-19 pandemic may reveal the importance of this tacit and "transparent" connectedness that we typically (before COVID) took it for granted.

[15]The other major classificatory system used in contemporary psychiatry is the ICD-10 (International Classification of Diseases, World Health Organisation). While there are some important differences between DSM and ICD, both largely agree upon the diagnostic criteria for DPD, which are the following: (a) persistent symptoms of DP/DR not occurring as part of another disorder or be directly substance-induced; (b) the individual should not be suffering from psychosis (which would imply a different diagnosis, such as schizophrenia). DSM adds the criterion (c) there should be significant distress and/or functional impairment. This seems appropriate, as otherwise it is hard to argue that the phenomena can usefully be seen as pathological (Medford et al., 2005).

[16]"Now I experience depersonalisation more of a lack of narrative; moments seem to melt away as soon as they have passed and life appears to pass as a series of unrelated frames. (...) You may be able to function, but creating a credible narrative for your life would be challenging" (Ciaunica and Charlton, 2018).

people feel "not fully real" (Simeon and Abugel, 2008; Sierra, 2009; Medford, 2012).

The prevalence of DPD is around 1–2% in the general population (Hunter et al., 2004), with onset typically occurring before age 25. Strikingly, feelings of depersonalization are the third most common psychological symptom reported in the general population (after anxiety and low mood), especially among young people (Simeon et al., 2003). Yet, the underlying neural and computational mechanisms as well as its phenomenological markers remain poorly understood (see Seth et al., 2011 for an early attempt).

The experience of a "split" between the self and the body–strikingly described as feeling trapped in one's head (mind) and outside one's body (Ciaunica et al., 2020)–is one of the most frequently cited symptoms in DPD (Sierra and David, 2011). Sierra (2009) lists four prominent types of anomalous body experiences in DPD: (1) lack of body ownership, (2) feelings of loss of agency, (3) disembodiment feelings, and (4) somatosensory distortions. Empirical evidence for this disrupted bodily sensory processing comes from studies that report disrupted physiological responses in patients with DPD, compared with healthy participants (Sierra et al., 2002; Owens et al., 2015; Dewe et al., 2018). DPD has also been linked to disrupted activity in neuronal regions underlying somatic processing (Medford et al., 2012; Lemche et al., 2013) and the vestibular system (Jáuregui Renaud, 2015), which is responsible for providing information about the body's position in space (Ferrè and Haggard, 2016).

A detailed discussion on this condition would lead to a substantial digression (see Simeon and Abugel, 2008; Sierra, 2009; Sierra and David, 2011; Medford, 2012; Billon, 2016; Ciaunica et al., 2020). Here, we propose to narrow our focus on relational aspects between the self and the world, dramatically reported by a patient as follows:

> "When the depersonalization is very deep, I still seek to 'be' someone else because it feels like that constant source of interaction is the only thing that allows me to maintain a connection with the world. I'll also seek physical contact with whoever I'm with. It almost feels as though I need to be that other person because my own sense of self is not strong enough in that moment to sustain me" (Ciaunica and Fotopoulou, 2017).

This observation may be crucial for potential interventions and therapies designed to make DPD people aware of what *still connects* them with their self, body, and the world, despite the fact that they feel dramatically alienated. Specifically, future work needs to disentangle and contrast the role of proximal vs. distal senses in shaping the experience of being a self-present in the world and fully connected with one's body.

If our argumentation throughout our paper is correct, then future research needs to assess whether active engagements with the world and others via proximal (tactile and multisensory) interactions enhance the sense of self, realness, and presence in people with DPD. By contrast, we speculate that distant visuospatial experiences–such as seeing/recognizing oneself in the mirror–will enhance feelings of estrangement and self-detachment, of being inside one's head. Moreover, we hypothesize that close and dynamic physical and synchronous interactions with their environment will make DPD people feel more present in their bodies, and less "trapped" in their minds. This is because, paradoxically, in order to get closer to one's self, one needs to be able to "forget" oneself (Ciaunica, 2020, submitted) and to be able to open to the world and others, via proximal tactile interactions.

The accent thus needs to be put on what connects us to ourselves and the reality, as opposed to what separates us from it, especially when we feel isolated and alienated. To use our metaphor, even if the transparent bridge that connects people with their self, body, and world is shattered, it is important to stress that it remains available and *can* be crossed. As Ratcliffe insightfully notes: "talk of feeling detached from body and world might best express an all-pervasive feeling of estrangement but, importantly, that feeling is *itself* a way of experiencing the body–world relationship and so one has not actually escaped from body and world at all" (2008, p. 131). We must thus use this fundamental connectedness to the world as a powerful tool to repair the "lost" connectedness to one's self.

CONCLUSION AND FUTURE DIRECTIONS

While traditionally perceptual experiences have been described as having a subject–object polarity, throughout this paper, we aimed at drawing attention at its relational aspect, i.e., what is happening *in between* these two poles, especially when the sensory signals are proximal as opposed to distal. Shifting focus from visuo-distal perception (easier to structure around a subject/object axis) to proximal perception (and particularly on touch) allowed us to outline the key relational role of tactile interactions with the physical and social environment. We motivated this shift in light of the new and highly influential PP theories of perception. If indeed whatever we perceive now is infused by whatever is perceived before, then it becomes crucial to endorse a dynamic and situated viewpoint in examining perception.

We argued that the term "next" here must include not only the temporal aspect (i.e., the processing of upcoming inputs) but also the spatial dimension (i.e., the processing of upcoming inputs next-to-my-body). This is because accurate information processing of proximal sensory signals is key to the organism's survival in the here and now.

We narrowed our focus on tactile proximal interactions that are close to our bodies. We explored the special relationship between touch and body, and argued in line with the phenomenological accounts, that tactile experiences form a pervasive yet transparent experiential background, typically unnoticed and taken for granted. The key claim was that the body–world relationship remains tacitly and ubiquitously in the background and the sense of self is inseparable from this relationship.

To support this claim, we reviewed developmental empirical work that points to the idea that (discriminative and affective) tactile inputs play a fundamental role in constituting perceptual experiences *in utero* and in early life. We then explored the idea that our transparent experiential background includes also garments (i.e., objects and materials close-to-our-skin) that can be viewed as a "second skin" or "extended skin." As such, they constitute what one may call, following

Clark and Chalmers (1998) and Gallagher (2005) seminal formulations, an "extended" body-image and body-schema.

Our speculative proposal stipulates that proximal and tactile perceptual engagements with the physical and social world may form a pervasive yet transparent experiential bridge, typically unnoticed and taken for granted. Keeping this "bridge" open is essential in constituting the feeling of being real, present, and immersed in the world, especially when one's sense of self is shattered, as in the case of Depersonalization Disorder. If this is so, then future work and interventions need to focus on initiating and restoring dynamic and proximal engagements with our social and physical world in order to rebuild the "invisible" transparent yet essential basis of our sense of self and sense of presence in the world.

In line with this approach, we have recently piloted the sense of connection through touch in a set of experiential interventions with the microphenomenologist (Petitmengin et al., 2018). We have explored the aspect of quality of movements in textile experience further with designers through the Micro-phenomenology interview method[17] (Petitmengin, 2006; Petitmengin et al., 2013, 2018) to go deeper into pre-verbal and non-verbal experiences (i.e., movement and touch behavior).

Briefly, participants explored with their hands and with closed eyes unknown objects for an extended time. They were then invited to share the experiences focusing on how they became aware of the interface of touch. Participants typically used images that were porous: "I felt fuzziness" often integrating social and spatial domains; "I felt open toward you." Some would spontaneously describe a difference between touching and being touched. This distinction was subsequently frontloaded (Gallagher, 2003) into an exploration of another object. Participants were again invited to share these experiences, particularly noticing the distinction between touching and being touched, and on how sense of self and agency was experienced[18]. Such simple intervention appears to have the potentials to make visible those processes that connect us to our surroundings in ways that are both proximal and intimate. Future research will explore this in the context of depersonalization and other disturbances in embodied self-perception.

AUTHOR CONTRIBUTIONS

AC wrote the first draft of the paper. BP contributed with the textile/second skin section and provided comments on the first draft. AR and AF provided feedback on the final draft of the manuscript. All authors contributed to the article and approved the submitted version.

FUNDING

AC was supported by. Bial Foundation grant n° 157/16. AF was supported by a ERC grant n° ERC- 2018-COG-818070. BP was supported by the Burberry Material Futures Research Group.

[17]This method combines psychology and phenomenology approaches to obtain a first-person, fine-grained verbal description of subjective experience that is generally inaccessible or difficult to articulate (Varela and Shear, 1999), and it reduces post-rationalization (Hogan et al., 2015). By using specific guidance (through questions and non-verbal cues), it helps people to become aware of nuances in the experience being described (Petitmengin et al., 2013). Hence, the method is presented as a "psychological microscope" (Bitbol and Petitmengin, 2013).

[18]A video detailing an excerpt of the experiment can be found here https://www.eer.info/activities/being-aware-sharpening-our-tools.

REFERENCES

Ainsworth, M. S. (1979). Infant–mother attachment. *Am. Psychol.* 34:932.

Apps, M. A., and Tsakiris, M. (2013). The free-energy self: a predictive coding account of self-recognition. *Neurosci. Biobehav. Rev.* 41, 85–97. doi: 10.1016/j.neubiorev.2013.01.029

Barnard, M. (Ed.). (2007). *Fashion Theory: A Reader*. London: Routledge.

Barwitch, A. (2020). *Smellosophy*. Oxford: Harvard University Press.

Bellieni, C. V., Cordelli, D. M., Marchi, S., Ceccarelli, S., Perrone, S., Maffei, M., et al. (2007). Sensorial saturation for neonatal analgesia. *Clin. J. Pain* 23, 219–221. doi: 10.1097/AJP.0b013e31802e3bd7

Billon, A. (2016).Making sense of the cotard syndrome: insights from the study of depersonalization. *Mind Langu.* 31, 356–391. doi: 10.1111/mila.12110

Bitbol, M., and Petitmengin, C. (2013). A defense of introspection from within. *Constr. Found.* 8, 269–279. Available online at: http://constructivist.info/8/3/269

Blanke, O., and Metzinger, T. (2009). Full-body illusions and minimal phenomenal selfhood. *Trends Cogn. Sci.* 13, 7–13. doi: 10.1016/j.tics.2008.10.003

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

Bremmer, F., Klam, F., Duhamel, J. R., Ben Hamed, S., and Graf, W. (2002). Visual–vestibular interactive responses in the macaque ventral intraparietal area (VIP). *Eur. J. Neurosci.* 16, 1569–1586. doi: 10.1046/j.1460-9568.2002.02206.x

Breward, C. (2003). *Fashion*. Oxford: Oxford University Press.

Brumfield, E. (2006). Cloth, gender, continuity and change: fabricating unity in anthropology. *Am. Anthropol.* 108, 862–877. doi: 10.1525/aa.2006.108.4.862

Bystrova, K. (2009). Novel mechanism of human fetal growth regulation: a potential role of lanugo, vernix caseosa and a second tactile system of unmyelinated low-threshold C-afferents. *Med. Hypoth.* 72, 143–146. doi: 10.1016/j.mehy.2008.09.033

Carruthers, P. (1996). *Language, Thoughts and Consciousness. An Essay in Philosophical Psychology*. Cambridge: Cambridge University Press.

Castiello, U., Becchio, C., Zoia, S., Nelini, C., Sartori, L., Blason, L., et al. (2010). Wired to be social: the ontogeny of human interaction. PLoS ONE 5:e13199. doi: 10.1371/journal.pone.0013199

Ciaunica, A. (2019). The 'meeting of bodies': empathy and basic forms of shared experiences. *Topoi* 38, 185–195. doi: 10.1007/s11245-017-9500-x

Ciaunica, A. (2020). Introduction: The Relational Self: Basic Forms of Self-Awareness. *Topoi* 39, 501–507. doi: 10.1007/s11245-020-09689-z

Ciaunica, A., and Charlton, J. (2018). *When the Self Slips: What Depersonalization Can Say About the Self*. Available online at: https://aeon.co/essays/what-can-depersonalisation-disorder-say-about-the-self

Ciaunica, A., Charlton, J., and Farmer, H. (2020). When the window cracks: transparency and the fractured self in depersonalisation. *Phenomenol. Cogn. Sci.* 20, 1–19. doi: 10.1007/s11097-020-09677-z

Ciaunica, A., Constant, A., Preissl, H., and Fotopoulou, A. (2021). The first prior: From co-embodiment to co-homeostasis in early life. *Conscious Cogn.* (2021) 91:103117. doi: 10.1016/j.concog.2021.103117

Ciaunica, A., and Crucianelli, L. (2019). Minimal self-awareness: from within a developmental perspective. *J. Conscious. Stud.* 26, 207–226.

Ciaunica, A., and Fotopoulou, A. (2017). "The touched self: psychological and philosophical perspectives on proximal intersubjectivity and the self," in *Embodiment, Enaction, and Culture—Investigating the Constitution of the Shared World*, eds C. Durt, T. Fuchs, and C. Tewes (Cambridge, MA: MIT Press).

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

Clark, A., and Chalmers, D. (1998). The extended mind. *Analysis* 58, 7–19.

Cléry, J., Guipponi, O., Wardak, C., and Hamed, S. B. (2015). Neuronal bases of peripersonal and extrapersonal spaces, their plasticity and their dynamics: knowns and unknowns. *Neuropsychologia* 70, 313–326. doi: 10.1016/j.neuropsychologia.2014.10.022

Colombetti, G. (2016). "Affective incorporation," in *Phenomenology for the Twenty-First Century. Palgrave Macmillan*, eds J. Simmons and J. Hackett (London), 231–249.

Conant, R. C., and Ross Ashby, W. (1970). Every good regulator of a system must be a model of that system. *Int. J. Syst. Sci.* 1, 89–97.

Craig, A. D. (2009). How do you feel-now? The anterior insula and human awareness. *Nat. Rev. Neurosci.* 10, 59–70. doi: 10.1038/nrn2555

Crane, D. (2012). *Fashion and Its Social Agendas: Class, Gender, and Identity in Clothing*. Chicago; London: University of Chicago Press.

Crane, T., and French, G. (2015). "The problem of perception," in *Stanford Encyclopaedia of Philosophy*, ed E. N. Zalta. Available online at: http://plato.stanford.edu/entries/perception-problem/

Damasio, A. (1994). *Descartes' Error: Emotion, Reason, and the Human Brain*. New York, NY: G.P. Putnam's Sons.

De Jaegher, H., and Di Paolo, E. A. (2007). Participatory sense-making: an enactive approach to social cognition. *Phenomenol. Cogn. Sci.* 6, 485–507. doi: 10.1007/s11097-007-9076-9

de Vignemont, F. (2018). Peripersonal perception in action. *Synthese*, 1–18. doi: 10.1007/s11229-018-01962-4

de Vignemont, F., and Iannetti, G. D. (2015). How many peripersonal spaces? *Neuropsychologia* 70, 327–334.

Delafield-Butt, J. T., and Gangopadhyay, N. (2013). Sensorimotor intentionality: the origins of intentionality in prospective agent action. *Dev. Rev.* 33, 399–425. doi: 10.1016/j.dr.2013.09.001

Dewe, H., Watson, D. G., Kessler, K., and Braithwaite, J. J. (2018). The depersonalized brain: new evidence supporting a distinction between depersonalization and derealization from discrete patterns of autonomic suppression observed in a non-clinical sample. *Conscious. Cogn.* 63, 29–46. doi: 10.1016/j.concog.2018.06.008

di Pellegrino, G., and Làdavas, E. (2015). Peripersonal space in the brain. *Neuropsychologia* 66, 126–133. doi: 10.1016/j.neuropsychologia.2014.11.011

di Pellegrino, G., Làdavas, E., and Farné, A. (1997). Seeing where your hands are. *Nature* 388:730.

Entwistle, J. (2000). *The Fashioned Body: Fashion. Dress and Modern Social Theory*. London: Wiley.

Faivre, N., Arzi, A., Lunghi, C., and Salomon, R. (2017). Consciousness is more than meets the eye: a call for a multisensory study of subjective experience. *Neurosci. Conscious.* 2017:nix003. doi: 10.1093/nc/nix003

Felipe, N. J., and Sommer, R. (2017). "Invasions of personal space," in *Sociological Methods* (London: Routledge), 458–468.

Ferrè, E. R., and Haggard, P. (2016). The vestibular body: vestibular contributions to bodily representations. *Cogn. Neuropsychol.* 33, 67–81. doi: 10.1080/02643294.2016.1168390

Field, T. (2001). *Touch*. Boston, MA: The MIT Press.

Fotopoulou, A. (2015). The virtual bodily self: mentalisation of the body as revealed in anosognosia for hemiplegia. *Conscious. Cogn.* 33, 500–510. doi: 10.1016/j.concog.2014.09.018

Fotopoulou, A., and Tsakiris, M. (2017). Mentalizing homeostasis: the social origins of interoceptive inference. *Neuropsychoanalysis* 19, 3–28. doi: 10.1080/15294145.2017.1294031

Friston, K. (2005). A theory of cortical responses. *Philos. Trans. R. Soc. Lond.* 360, 815–36. doi: 10.1098/rstb.2005.1622

Friston, K. (2008). Hierarchical models in the brain. *PLoS Comput. Biol.* 4:11. doi: 10.1371/journal.pcbi.1000211

Friston, K. J., and Stephan, K. E. (2007). Free-energy and the brain. *Synthese* 159, 417–458. doi: 10.1007/s11229-007-9237-y

Fuchs, T. (2005). Corporealized and disembodied minds: a phenomenological view of the body in melancholia and schizophrenia. *Philos. Psychiatry Psychol.* 12, 95–107. Available online at: https://psycnet.apa.org/record/2005-15458-001

Fulkerson, M. (2014). *The First Sense: A Philosophical Study of Human Touch*. Cambridge, MA: MIT Press.

Gallace, A., and Spence, C. (2010). The science of interpersonal touch: an overview. *Neurosci. Biobehav. Rev.* 34, 246–259. doi: 10.1016/j.neubiorev.2008.10.004

Gallagher, S. (1986). Body image and body schema: a conceptual clarification. *J. Mind Behav.* 7, 541–554.

Gallagher, S. (2000). Philosophical conceptions of the self: implications for cognitive science. *Trends Cogn. Sci.* 4, 14–21. doi: 10.1016/S1364-6613(99)01417-5

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

Gallagher, S. (2005). *How the Body Shapes the Mind*. Oxford: Oxford University Press.

Gibson, J. J. (1979). *The Ecological Approach to Visual Perception*. Hillsdale, NJ: LEA.

Gillath, O., Bahns, A. J., Ge, F., and Crandall, C. S. (2012). Shoes as a source of first impressions. *J. Res. Personal.* 46, 423–430. doi: 10.1016/j.jrp.2012.04.003)

Gorman, W. (1969). *Body Image and the Image of the Brain*. St Louis, MO: Warren H Green.

Guy, A., Green, E., and Banim, M. (Eds.). (2001). *Through the Wardrobe: Women's Relationships with their Clothes*. Oxford: Berg.

Hall, E. T. (1990). *Understanding Cultural Differences*. Boston, MA: Intercultural Press.

Hall, E. T., Birdwhistell, R. L., Bock, B., Bohannan, P., Diebold, A. R. Jr., Durbin, M., et al. (1968). Proxemics [and comments and replies]. *Curr. Anthropol.* 9, 83–108.

Hannover, B., and Kühnen, U. (2002). "The clothing makes the self" via knowledge activation. *J. Appl. Soc. Psychol.* 32, 2513–2525. doi: 10.1111/j.1559-1816.2002.tb02754.x

Hansen, K. T. (2004). The world in dress: anthropological perspectives on clothing, fashion and culture. *Ann. Rev. Anthropol.* 33, 369–392. doi: 10.1146/annurev.anthro.33.070203.143805

Harman, G. (1990). The intrinsic quality of experience. *Philos. Perspect.* 4, 31–52. doi: 10.2307/2214186

Hata, T. (2016). Current status of fetal neurodevelopmental assessment: four-dimensional ultrasound study. *J. Obstetr. Gynaecol. Res.* 42, 1211–1221. doi: 10.1111/jog.13099

Heed, T., Habets, B., Sebanz, N., and Knoblich, G. (2010). Others' actions reduce crossmodal integration in peripersonal space. *Curr. Biol.* 20, 1345–1349. doi: 10.1016/j.cub.2010.05.068

Henderson, C. (1990). Cloth and Human Experience. *Science* 249, 1588–1590.

Hogan, T., Hinrichs, U., and Hornecker, E. (2015). The elicitation interview technique: Capturing people's experiences of data representations. *IEEE Trans. Visual. Comput. Graph.* 22, 2579–2593. doi: 10.1109/TVCG.2015.2511718

Hohwy, J. (2013). *The Predictive Mind*. Oxford: Oxford University Press.

Hohwy, J. (2014). The self-evidencing brain. *Noûs* 50, 259–285. doi: 10.1111/nous.12062

Hohwy, J. (2020). New directions in predictive processing. *Mind Lang.* 35, 209–223. doi: 10.1111/mila.12281

Hohwy, J., and Michael, J. (2017). "Why should any body have a self?," in *Self-Consciousness and the Body*. eds de Vignemont F. and Alsmith AJ. (Cambridge MA: MIT Press), 363–391

Hunter, E. C., Sierra, M., and David, A. S. (2004). The epidemiology of depersonalization and derealisation. a systematic review. *Soc. Psychiatry Psychiatr. Epidemiol.* 39, 9–18. doi: 10.1007/s00127-004-0701-4

Husserl, E. (1989). *Ideas Pertaining to a Pure Phenomenology and to a Phenomenological Philosophy: Second Book Studies in the Phenomenology of Constitution, Vol. 3*. Amsterdam: Springer Science & Business Media.

James, W. (1890). The perception of reality. *Principles Psychol.* 2, 283–324.

Jáuregui Renaud, K. (2015). Vestibular function and depersonalization/derealization symptoms. *Multisensory Res.* 28, 637–651. doi: 10.1163/22134808-00002480

Kirchhoff, M. D., and Kiverstein, J. (2019). *Extended Consciousness and Predictive Processing: A Third Wave View*. Routledge.

Kirsch, L. P., Besharati, S., Papadaki, C., Crucianelli, L., Bertagnoli, S., Ward, N., et al. (2020). Damage to the right insula disrupts the perception of affective touch. *Elife* 9:e47895. doi: 10.7554/eLife.47895

Kirsch, L. P., Krahé, C., Blom, N., Crucianelli, L., Moro, V., Jenkinson, P. M., et al. (2018). Reading the mind in the touch: Neurophysiological specificity in the communication of emotions by touch. *Neuropsychologia* 116, 136–149. doi: 10.1016/j.neuropsychologia.2017.05.024

Krahé, C., Drabek, M. M., Paloyelis, Y., and Fotopoulou, A. (2016). Affective touch and attachment style modulate pain: a laser-evoked potentials study. *Philos. Trans. R. Soc. B Biol. Sci.* 371:20160009. doi: 10.1098/rstb.2016.0009

Làdavas, E. (2002). Functional and dynamic properties of visual peripersonal space. *Trends Cogn. Sci.* 6, 17–22. doi: 10.1016/s1364-6613(00)01814-3

Legrand, D. (2006). The bodily self: The sensorimotor roots of pre-reflective self-consciousness. *Phenomenol. Cogn. Sci.* 5, 89–118. doi: 10.1007/s11097-005-9015-6

Lemche, E., Brammer, M. J., David, A. S., Surguladze, S. A., Phillips, M. L., Sierra, M., et al. (2013). Interoceptive-reflective regions differentiate alexithymia traits in depersonalization disorder. *Psychiatry Res. Neuroimaging* 214, 66–72. doi: 10.1016/j.pscychresns.2013.05.006

Limanowski, J., and Blankenburg, F. (2013). Minimal self-models and the free-energy principle. *Front. Hum. Neurosci.* 7:547. doi: 10.3389/fnhum.2013.00547

Limanowski, J., and Friston, K. (2018). 'Seeing the dark': grounding phenomenal transparency and opacity in precision estimation for active inference. *Front. Psychol.* 9:643. doi: 10.3389/fpsyg.2018.00643

Limanowski, J., and Friston, K. (2020). Attentional modulation of vision versus proprioception during action. *Cereb. Cortex* 30, 1637–1648. doi: 10.1093/cercor/bhz192

Lloyd, D. M., Gillis, V., Lewis, E., Farrell, M. J., and Morrison, I. (2013). Pleasant touch moderates the subjective but not objective aspects of body perception. *Front. Behav. Neurosci.* 7:207. doi: 10.3389/fnbeh.2013.00207

Löken, L. S., Wessberg, J., Morrison, I., McGlone, F., and Olausson, H. (2009). Coding of pleasant touch by unmyelinated afferents in humans. *Nat. Neurosci.* 12, 547–548. doi: 10.1038/nn.2312

Lurie, A. (2000). *The Language of Clothes*. New York, NY: Henry Holt and Company.

Maravita, A., Spence, C., and Driver, J. (2003). Multisensory integration and the body schema: close to hand and within reach. *Curr. Biol.* 13, R531–R539. doi: 10.1016/S0960-9822(03)00449-4

Maravita, A., Spence, C., Sergent, C., and Driver, J. (2002). Seeing your own touched hands in a mirror modulates cross-modal interactions. *Psychol. Sci.* 13, 350–355. doi: 10.1111/j.0956-7976.2002.00463.x

Martin, M. G. (1992). Perception, concepts, and memory. *Philos. Rev.* 101, 745–763.

Martin, M. G. F. (1993). Sense modalities and spatial properties. *Spat. Represent.* 206–218.

Martin, M. G. F. (1995). "Bodily awareness: a sense of ownership," in *The Body and the Self,* eds J. L. Bermúdez, T. Marcel, and N. Eilan (Cambridge, MA: MIT Press).

Marx, V., and Nagy, E. (2015). Fetal behavioural responses to maternal voice and touch. PLoS ONE 10:e0129118. doi: 10.1371/journal.pone.0129118

McGlone, F., Vallbo, Å., Olausson, H., Löken, L., and Wessberg, J. (2007). Discriminative touch and emotional touch. *Can. J. Exp. Psychol.* 61, 173–183. doi: 10.1037/cjep2007019

McGlone, F., Wessberg, J., and Olausson, H. (2014). Discriminative and affective touch: sensing and feeling. *Neuron* 82, 737–755. doi: 10.1016/j.neuron.2014.05.001

Medford, N. (2012). Emotion and the unreal self: depersonalization disorder and de-affectualization. *Emot. Rev.* 4, 139–144. doi: 10.1177/1754073911430135

Medford, N., Sierra, M., Baker, D., and David, A. (2005). Understanding and treating depersonalisation disorder. *Adv. Psychiatric Treatment* 11, 92–100. doi: 10.1192/apt.11.2.92

Medford, N., Sierra, M., Stringaris, A., Giampietro, V., Brammer, M. J., and David, A. S. (2012). Emotional experience and awareness of self: functional MRI studies of depersonalization disorder. Front. Psychol. 7:432. doi: 10.3389/fpsyg.2016.00432

Merleau-Ponty, M. (1962). Un inédit de Maurice Merleau-Ponty. *Revue de Métaphysique et de Morale* 67, 401–409.

Metzinger, T. (2003). Phenomenal transparency and cognitive self-reference. *Phenomenol. Cogn. Sci.* 2, 353–393. doi: 10.1023/B:PHEN.0000007366.42918.eb

Mida, I., and Kim, A. (2015). *The Dress Detective: A Practical Guide to Object-Based Research in Fashion*. London: Bloomsbury Publishing.

Montagu, A. (1971). *Touching: The Human Significance of the Skin*. New York, NY: Columbia University Press.

Moore, G. E. (1903). The refutation of idealism. *Mind* 12, 433–453. doi: 10.1093/mind/XII.4.433

Mori, H., and Kuniyoshi, Y. (2010). "A human fetus development simulation: Self-organization of behaviors through tactile sensation," in *2010 IEEE 9th International Conference on Development and Learning,* (Ann Arbor, MI: IEEE), 82–87.

Morrison, I., Loken, L. S., and Olausson, H. (2010). The skin as a social organ. *Exp. Brain Res.* 204, 305–314. doi: 10.1007/s00221-009-2007-y

Noë, A. (2004). *Action in Perception*. Cambridge, MA: MIT press.

O'Connor, K. (2010). How smart is smart? T-shirts, wellness, and the way people feel about "medical" textiles. *Textile* 8, 50–66. doi: 10.2752/175183510X12580391270029

Owens, A. P., David, A. S., Low, D. A., Mathias, C. J., and Sierra-Siegert, M. (2015). Abnormal cardiovascular sympathetic and parasympathetic responses to physical and emotional stimuli in depersonalization disorder. *Front. Neurosci.* 9:89. doi: 10.3389/fnins.2015.00089

Panagiotopoulou, E., Filippetti, M. L., Tsakiris, M., and Fotopoulou, A. (2017). Affective touch enhances self-face recognition during multisensory integration. *Sci. Rep.* 7:12883. doi: 10.1038/s41598-017-13345-9

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

Parr, T., Rees, G., Friston, K. J. (2018). Computational neuropsychology and Bayesian inference. Front. Hum. Neurosci. 12:61. doi: 10.3389/fnhum.2018.00061

Peluchette, J. V., and Karl, K. (2007). The impact of workplace attire on employee self-perceptions. *Hum. Res. Develop. Q.* 18, 345–360. doi: 10.1002/hrdq.1208

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., Remillieux, A., Cahour, B., and Carter-Thomas, S. (2013). A gap in Nisbett and Wilson's findings? A first-person access to our cognitive processes. *Conscious. Cognit.* 22, 654–669. doi: 10.1016/j.concog.2013.02.004

Petitmengin, C., Remillieux, A., and Valenzuela-Moguillansky, C. (2018). Discovering the structures of lived experience. Towards a micro-phenomenological analysis method. *Phenomenol. Cogn. Sci.* 18, 691–730. doi: 10.1007/s11097-018-9597-4

Petreca, B., Bianchi-Berthouze, N., and Baurley, S. (2015). "How do designers feel textiles?" in *Proceedings of the International Conference on Affective Computing and Intelligent Interaction (ACII'15). Xi'an, China, 21-24 September 2015* (Washington, DC: IEEE), 982–987. doi: 10.1109/ACII.2015.7344695

Piontelli, A. (2010). *Development of Normal Fetal Movements: The First 25 Weeks of Gestation*. Milan: Springer. doi: 10.1007/978-88-470-1402-2

Piontelli, A. (2014). *Development of Normal Fetal Movements*. Amsterdam: Springer. doi: 10.1007/978-88-470-5373-1

Qin, P., Wang, M., and Northoff, G. (2020). Linking bodily, environmental and mental states in the self–a three-level model based on a meta-analysis. *Neurosci. Biobehav. Rev.* 115, 77–95. doi: 10.1016/j.neubiorev.2020.05.004

Quintero, A. M., and de Jaegher, H. (2020). Pregnant agencies: movement and participation in maternal–fetal interactions. *Front. Psychol.* 11:1977. doi: 10.3389/fpsyg.2020.01977

Ratcliffe, M. (2008). *Feelings of Being: Phenomenology, Psychiatry and the Sense of Reality*. Oxford: Oxford University Press. doi: 10.1093/med/9780199206469.001.0001

Ratcliffe, M. (2013). Touch and the sense of reality," in *The Hand, and Organ of the Mind* ed Radman, Z. 131–157.

Rizzolatti, G., Scandolara, C., Matelli, M., and Gentilucci, M. (1981). Afferent properties of periarcuate neurons in macaque monkeys. I. Somatosensory responses. *Behav. Brain Res.* 2, 125–146.

Salgado-Montejo, A., Moreno, A. M., Salgado, R., Greiffenstein, D., and Valencia, J. A. (2018). The *Magical Body Mirror: Linking Personality, Clothing Preference and Nonverbal Cues to Understand Decision-Making.* Fusion 2018: Big Data World. Amsterdam: ESOMAR 2018.

Sass, L.A., and Parnas, J. (2003). Schizophrenia, consciousness, and the self. *Schizophrenia Bull.* 29, 427–444. doi: 10.1093/oxfordjournals.schbul.a007017

Schilbach, L., Timmermans, B., Reddy, V., Costall, A., Bente, G., and Schlicht, T. (2013). Toward a second-person neuroscience. *Behav. Brain Sci.* 36, 393–414. doi: 10.1017/S0140525X12000660

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

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. K., Suzuki, K., and Critchley, H. D. (2011). An interoceptive predictive coding model of conscious presence. *Front. Psychol.* 2:395. doi: 10.3389/fpsyg.2011.00395

Seth, A. K., and Tsakiris, M. (2018). Being a beast machine: the somatic basis of selfhood. *Trends Cogn. Sci.* 22, 969–981. doi: 10.1016/j.tics.2018.08.008

Sierra, M. (2009). *Depersonalisation: A New Look at a Neglected Syndrome*. Cambridge: Cambridge University Press. doi: 10.1017/CBO9780511730023

Sierra, M., and Berrios, G. E. (1998). Depersonalization: neurobiological perspectives. *Biol. Psychiatry* 44, 898–908. doi: 10.1016/S0006-3223(98)00015-8

Sierra, M., and David, A. S. (2011). Depersonalization: a selective impairment of self-awareness. *Conscious. Cogn.* 20, 99–108. doi: 10.1016/j.concog.2010.10.018

Sierra, M., Senior, C., Dalton, J., McDonough, M., Bond, A., Phillips, M. L., et al. (2002). Autonomic response in depersonalization disorder. *Archiv. Gen. Psychiatry* 59, 833–838. doi: 10.1001/archpsyc.59.9.833

Simeon, D., and Abugel, J. (2008). *Feeling Unreal: Depersonalization Disorder and the Loss of the Self*. Oxford; New York, NY: Oxford University Press.

Simeon, D., Knutelska, M., Nelson, D., and Guralnik, O. (2003). Feeling unreal: a depersonalization disorder update of 117 cases. *J. Clin. Psychiatry* 64, 990–997. doi: 10.4088/JCP.v64n0903

Slepian, M. L., Ferber, S. N., Gold, J. M., and Rutchick, A. M. (2015). The cognitive consequences of formal clothing. *Soc. Psychol. Personal. Sci.* 6, 661–668. doi: 10.1177/1948550615579462

Sommer, R. (1969). *Personal Space. The Behavioral Basis of Design.* ed Cliffs, E. (NJ: Prentice-Hall, Inc).

Stack, D. M., and Muir, D. W. (1992). Adult tactile stimulation during face-to-face interactions modulates five-month-olds' affect and attention. *Child Dev.* 63, 1509–1525. doi: 10.2307/1131572

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

Thompson, E. (2007). Mind in Life: Biology, Phenomenology, and the Sciences of Mind. Cambridge, MA: Harvard University Press.

Twigg, J., and Buse, C. E. (2013). Dress, dementia and the embodiment of identity. *Dementia* 12, 326–336. doi: 10.1177/1471301213476504

Tye, M. (1999). Phenomenal consciousness: the explanatory gap as a cognitive illusion. *Mind* 108, 705–725. doi: 10.1093/mind/108.432.705

van Stralen, H. E., van Zandvoort, M. J., Hoppenbrouwers, S. S., Vissers, L. M., Kappelle, L. J., and Dijkerman, H. C. (2014). Affective touch modulates the rubber hand illusion. *Cognition* 131, 147–158. doi: 10.1016/j.cognition.2013.11.020

Varela, F., Thompson, E., and Rosch, E. (1991). The Embodied Mind: Cognitive Science and Human Experience. Cambridge, MA: MIT Press. doi: 10.7551/mitpress/6730.001.0001

Varela, F. J., and Shear, J. (1999). First-person methodologies: What, why, how. *JCS* 6, 1–14.

von Mohr, M., Kirsch, L. P., and Fotopoulou, A. (2017). The soothing function of touch: affective touch reduces feelings of social exclusion. *Sci. Rep.* 7:13516. doi: 10.1038/s41598-017-13355-7

von Mohr, M., Krahé, C., Beck, B., and Fotopoulou, A. (2018). The social buffering of pain by affective touch: a laser-evoked potential study in romantic couples. *Soc. Cogn. Affect. Neurosci.* 13, 1121–1130. doi: 10.1093/scan/nsy085

Weber, S., and Mitchell, C. (eds.). (2004). *Not Just Any Dress: Narratives of Memory, Body, and Identity,* Vol. 220. New York, NY: Peter Lang.

Weiss, S. J., Wilson, P., Hertenstein, M. J., and Campos, R. (2000). The tactile context of a mother's caregiving: implications for attachment of low birth weight infants. *Infant Behav. Dev.* 23, 91–111. doi: 10.1016/S0163-6383(00)00030-8

Zoia, S., Blason, L., D'Ottavio, G., Bulgheroni, M., Pezzetta, E., Scabar, A., et al. (2007). Evidence of early development of action planning in the human foetus: a kinematic study. *Exp. Brain Res.* 176, 217–226. doi: 10.1007/s00221-006-0607-3

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

CONCEPTUAL ANALYSIS
published: 09 September 2021
doi: 10.3389/fpsyg.2021.539555

Bodies in the Novel *Infinite Jest*

Ana Chapman[1]*, *Silvia Chapman*[2] and *Stephanie Cosentino*[2,3]

[1] Department of English, French and German, Faculty of Arts, University of Málaga, Málaga, Spain, [2] Cognitive Neuroscience Division, Taub Institute for Research on Alzheimer's Disease and Gertrude H. Sergievsky Center, Columbia University Irving Medical Center, New York, NY, United States, [3] Department of Neurology, Columbia University Irving Medical Center, New York, NY, United States

This manuscript provides a literary analysis of the use of bodies in the novel *Infinite Jest* by David Foster Wallace. The novel describes a world where oversaturation of external stimulation leads to the perception of mind and body of self of an individual as prosthetic parts, malleable and deformed, wherein the mind fails to feel bodily sensations and characters experience a complete disconnectedness from the self and others. Indeed, the disembodiment of characters and sensations of disconnection leads them to a compulsive quest for connectedness through the use of masks, made-up feelings, mind–body hybrid pain, corporeal malleability, and prostheses. These portrayals of the disordered and disconnectedness between body and mind or self will be described and compared to clinical conditions characterized by a disconnection between mind and body and impaired body self-awareness. Through this exercise, we argue that the use of scientifically inspired pathologized bodies is a means of conveying the stance of Wallace on or criticism of the degradation of society through excessive entertainment.

Keywords: literature, David Foster Wallace, body self-awareness, embodiment, cognitive science

OPEN ACCESS

Edited by:
Anna Ciaunica,
University of Porto, Portugal

Reviewed by:
Simona Raimo,
University of Campania Luigi Vanvitelli, Italy
Elise Nykänen,
University of Helsinki, Finland

***Correspondence:**
Ana Chapman
achapman@uma.es

Specialty section:
This article was submitted to Theoretical and Philosophical Psychology, a section of the journal Frontiers in Psychology

Received: *01 March 2020*
Accepted: *09 August 2021*
Published: *09 September 2021*

Citation:
Chapman A, Chapman S and Cosentino S (2021) Bodies in the Novel Infinite Jest. Front. Psychol. 12:539555. doi: 10.3389/fpsyg.2021.539555

INTRODUCTION

Body and mind, and their relationship, have long been the object of interest across scientific, social, literary, and artistic fields. Descriptions, depictions, and interpretations of body and mind, in close connection to ownership and embodiment, are continuously evolving concepts representing a blending of ideas, figurative forms, and material conditions. Our lived experience is both nature- and culture-driven, giving rise to biological theories that include an analysis of social influences and by-products. The body has become an artifact increasingly represented as a double site of external and internal struggle. We argue that it is through an interdisciplinary perspective that we as humans attempt to understand the body as either the state of the self or an artistic or philosophical representation of an external force (i.e., society, divinity, etc.).

Throughout history, many corporeal representations have co-existed across philosophical, literary, and scientific perspectives. Images range from cave paintings wherein bodies were represented as simple mud figures focusing on action rather than form, to Haraway's (1991) cyborg or posthuman body, a hybrid creature composed of organisms and machines. Descriptions of the body in art have provided unique windows through which to view the state of philosophical and scientific quests that society undertakes to understand the role and purpose of the human body. Particularly, within literary work, the body and its connection to the mind have also been exploited as a medium to understand or express the social concerns of specific periods and cultures. Philosophers and artists, among others, have engaged in the discovery of the nature of the body while simultaneously incorporating knowledge and conceptions about the body derived from science, society, and religion (Fletcher, 1816; Shakespeare, 1980; Homer., 1987, 2003; Burton, 2001; Chaucer, 2003; Johnson, 2004; Wordsworth and Coleridge, 2013; Shelley, 2018

and in Dickens' work: Kryger, 2012). Using Illustration works depicting feminism, racism, and disability have also used the body as a means to represent these fights. As Robb and Harris (2013) state, the body enables us to "examine how historical circumstances create particular kinds of bodies, but [also] how particular kinds of bodies generate certain forms of history" (p. 4). From a scientific perspective, with the turn of the twentieth century, disciplines, such as neurology, neuroscience, neuropsychology, and psychology, flourished, and new theories of mind and body emerged. Diseases of the body were now understood through multifactorial lenses of neural underpinnings and cognitive and psychological processes. Common to all these disciplines is the newly found focus on the self and its disruption in the context of certain diseases.

This study aims to use the novel *Infinite Jest,* a twentieth-century novel wherein descriptions of bodily disturbances play a key part, to exemplify the intersectionality of artistic and scientific bodily perspectives reflective of the social and medical scientific context of the times. Although there have been many literary analyses of *Infinite Jest* such as those focusing on the themes of art vs. entertainment (Cioffi, 2000), addiction (Morris, 2001), solipsism and narcissism (Holland, 2013), and prosthetic description (Russell, 2010), among others, this study will focus solely on how bodies and body–mind connections are described in the novel. We suggest that contemporary (neuro)psychological and cognitive scientific works pertaining to body, mind, and self can be applied to understand the work of Wallace, and may even have served as sources of inspiration. Specifically, we propose that the physical distortions described by Wallace bring to mind contemporaneous neuroscientific models of disturbed disconnections between bodies and mind and body awareness, and in particular, how distortions in self-perception give rise to unique clinical phenomena. Descriptions of characters will be linked to theoretical and clinical constructs of embodiment (i.e., the "felt and conceptualized self" within own body of oneself; Crespi and Dinsdale, 2019), body image (i.e., the mental image or representation that individuals hold about their body shape, size, and form and the feelings associated with these; Slade, 1988), sense of ownership (i.e., the feeling that our body belongs to us; Gallagher, 2000), and sense of agency (e.g., the experience of being the agent that starts and controls an action; Moore and Fletcher, 2012).

This analysis widens the scope of how literature is a product of the mind and thus an interpretative device, rather than a verification of scientific theories of the mind; what literature aims to do, and what we attempt in this study, is not to "force" neuroscientific labels, but to examine how literary works can shed light on scientific perspectives that occupy a particular moment in history. This precisely connects to the field of cognitive literary studies that among other aspects analyze representations in literature from a cognitive lens (Hart, 2001; Spolsky, 2003; Richardson, 2010; Zunshine, 2010; Richardson and Spolsky, 2017). Particularly, we use a cognitive literary approach to analyze concepts of emotions, experience, and thoughts that are embedded, depicted, or encoded in literary works. Connections are made to embodied cognition (or the phenomenological and enactive streams of thought) where self-consciousness (of fictionalized characters, accounts of the author, or through the act of reading) arises from the living body. We will argue that through the intersectionality of clinical pathologies of the bodies within fictional characters, Wallace masterfully conveys his criticisms of the maladies of our current society, which seems to be becoming a normalized stance in society, abnormal embodiment and self.

INFINITE JEST: BACKGROUND OF THE NOVEL

Wallace is best known for his behemoth multilayered novel *Infinite Jest* and his nonfiction contributions. Published in 1996, *Infinite Jest* is a dense and complex 1,079-page novel with 388 endnotes depicting a dystopia originated from the overuse/overexposure and totalitarianism of the entertainment world in a country that resembles America.

Set in a near-future Boston, the novel examines the three main stories that crisscross and merge in many parts of the narrative. The three stories represent different sides of the entertainment world in which sport, addiction, and the media culture are analyzed in each intertwined narrative. First, tennis, one of the interests of Wallace, is represented through the Enfield Tennis Academy (E.T.A.) set up by Jim Incandenza, father of Orin, Hal, and Mario. Second, the main representation of addiction in the novel comes from the narration of the members of Ennet Alcohol and Drugs Rehabilitation House. This rehabilitation house establishes specific patterns of recovery through the commitment to a higher power. Third, the media is portrayed through a search of the *Assassins des Fauteuils Rollents* for Entertainment, a film, alternatively called Infinite Jest. The film, created by James Incandenza, is the ultimate stimulation that induces a catatonic death in its viewers. This film becomes the finest example of Wallace of how entertainment hinders action, willed choice, and experience through the destruction of body function, control, and sensations (basic bodily needs) to the extreme of possible death.

Infinite Jest is a book that works on obscurity and annularity, which results in an apparent non-concluding story of endless addiction and its effects. At the end of the novel, readers are left without much certitude of the course of events. The condition of Hal remains a mystery reflected in his unsettling smiling face and the uncertainty of his ability to speak as well as the condition of Don Gately at the hospital. Finally, the third story of *The Assassins,* the Canadian terrorists, ends with them on their way to Enfield Academy in search of deadly entertainment. Further, the novel is "fixated by its tools for chemical and electronic self-gratification [as it also] seems more prescient with the rollout of every new compulsively entertaining digital device" (Kalfus, 2010). It is a work that includes concerns and themes such as solipsism, language, terrorism, media, sport, addiction, self-reflexivity, irony, and politics. All events and characters intertwine in the discovery of what seems to reflect the loneliness sensed in each narrative, with crises of identity and self in an externally controlled and mediated world. Wallace points toward a "normality" through the inclusion of body abnormalities in

the novel, wherein bodies are represented as fragmented and disconnected from an individual oneself and lacking a sense of agency. This affects the perception of the reader and it is seemingly used by the author as a tool to illustrate how society is leading toward this figurative mental illness as the "normalized" stance.

REFLECTIONS ON NEUROSCIENCE IN *INFINITE JEST*

With the "cognitive turn" beginning in the 1990s, we can find a corresponding connection between literary works such as *Infinite Jest* with neuroscientific research. Cognitive interpretations of literary texts have a wide range of practice, values and perspectives, and have brought much discussion of the connection between both literature and neuroscience (e.g., effect on mind of the readers, interpretation of the text, etc). Caracciolo (2016) considers that a literary work "can serve as application or illustration of a scientific theory" but "evidence [cannot be found] for a particular view of the mind in a literary text, because textual evidence is different from scientific evidence" (pp. 192–193). That being said, fictional works, in general, then must be considered as what they are, literary productions that attempt to analyze and apply different concepts for understanding particular aspects of society. Furthermore, literary criticism allows a diversity of interpretations which, in this particular case, are based on the emerging cognitive literary approach surrounding theories of the mind where experience and consciousness seem to arise from bodily sensations, emotions, and its dialectic relationship with its surroundings.

Our study intends to illustrate neuroscientific theories of the time. Indeed, Wallace seems to have held a major concern regarding how body and mind disorders, as a result of entertainment and current societal trends, affect the identity of the self (full catalog of his library is available at Harry Ransom Center) and has been described as having a "career-long fascination with consciousness" (Burn, 2012, p. 373). Wallace himself struggled to understand his mental states (depression, addiction, and suicidal attempts). It is publicly known that he was admitted for substance addiction in 1989 in the psychiatric hospital associated with Harvard University. His knowledge of medicine is displayed in the medical terms and chemical components of drugs described in the novel. The abundance of medical and pharmaceutical language together with the many instances of mental illnesses such as the following, become a window to observe neuroscientific knowledge at his time.

> And so but since the old CBC documentary's thesis was turning out pretty clearly to be SCHIZOPHRENIA: BODY, the voiceover evinced great clipped good cheer as it explained that well, yes, poor old Fenton here was more or less hopeless as an extra-institutional functioning unit, but that, on the up-side, science could at least give his existence some sort of meaning by studying him very carefully to help learn how schizophrenia manifested itself in the human body's brain... that, in other words, with the aid of cutting-edge Positron-Emission Topography or "P.E.T." technology (since supplanted wholly by Invasive Digitals, Orin hears the developmental psychology graduate student mutter to herself, watching rapt over her cup, unaware that Orin's paralytically awake), they could scan and study how different parts of poor old Fenton's dysfunctional brain emitted positrons in a whole different topography than your average hale and hearty nondelusional God-fearing Alber-tan's brain, advancing science by injecting test-subject Fenton here with a special blood-brain-barrier-penetrating radioactive dye and then sticking him in the rotating body-sized receptacle of a P.E.T. Scanner. (Wallace, 2006, p. 47)

It must be noted that Wallace was obsessed with language, and he carefully explored the use of the most specific and accurate wording possible. Furthermore, his work was published during the "Decade of the Brain" and the "cognitive turn" in literature, leading to a flourish of cognitive literary studies that bring the science of the mind into literature (Jackson, 2000, 2003; Crane, 2001; Hart, 2001; Zunshine, 2010, 2017; Caracciolo, 2016; Burke and Troscianko, 2017). Bearing this in mind, we can present an analysis of the body in the novel in connection to contemporary studies on body and mind as well as to highlight knowledge of Wallace about theories of the mind.

DISCONNECTIONS IN BODY AND MIND

In *Infinite Jest*, the hedonistic search for pleasure in the overexposed mediated world influences lived experiences of the characters, which, we propose, are seen through disconnections of the physical body, as well as disconnections between physical body and mind.

Bodies that are incomplete or deformed pervade *Infinite Jest*. Descriptions of bodies are narrated through the depiction of separate, prosthetic, malleable, monstrous bodies challenging what one may consider the natural or classical body. Characters in the novel are represented as lopsided collections of parts or dissembled bodies, far from being coherent and unitary entities, they are anti-biological representations of lived bodies that will never reach satisfactory completion. Body parts are detached in the way Deleuze and Guattari (2004) define fractals, "aggregates whose number of dimensions is fractional rather than whole, or else with continuous variations" (p. 534). On the very first page, Hal identifies the people on the board as not being complete, akin to non-united bodies as if he were "surrounded by heads and bodies" (Wallace, 2006, p. 3). Wallace consciously focuses on limbs to produce an autonomous and disconnected sense of corporeality. Following analysis of Russell, limbs present exaggerated forms as a "tendency towards unequal distribution and "jagged seams" (Russell, 2010, p. 149). Tennis players are described as an assemble of fragmented body parts, as "most of the E.T.A. upperclassmen have these vivid shoe-and-shirt tans that give them the classic look of bodies hastily assembled from different parts of bodies, especially when you throw in the heavily muscled legs and usually shallow chests and the two arms of different sizes" (Wallace, 2006, p. 100). Even though exaggerated forms in sports can be considered normal in that context, continuous illustration of the novel of body disruptions sets the focus on these disruptions and possible

effects on the self. Don Gately, an ex-oral-narcotics addict, has "a massive and almost perfectly square head" (p. 55). Charles Tavis is somewhat impaired because "the two sides of his face didn't quite go together" (p. 521). Joelle Van Dyne, also called Madame Psychosis, belongs to the Union of the Hideously and Improbably Deformed (UHID) and calls for all "the prosthetically malmatched … the convulsively Tourettic. The Parkinsonianly tremulous… The teratoid of overall visage… The in any way asymmetrical" (p. 192) on the radio every night. Detached or separate parts are also presented in dreams of Orin in which:

> In the dream, it's understandably vital to Orin that he disengages his head from the phylacteryish bind of his mother's disembodied head, and he cannot. Last night's Subject's note indicates that at some point last night Orin had clutched her head with both hands and tried to sort of stiff-arm her, though not in an ungentle or complaining way (the note, not the stiff-arm). The apparent amputation of the Moms' head from the rest of the Moms appears in the dream to be clean and surgically neat: there is no evidence of a stump or any kind of nubbin of the neck, even, and it is as if the base of the round pretty head had been sealed, and also sort of rounded off, so that her head is a large living ball, a globe with a face, attached to his own head's face. (p. 46)

Through the creative process of Wallace, disturbances are also seen in the disconnected experiences of body and mind, in which individuals are unable to register or experience that they are in control of their actions (Fink et al., 1999). It is important to note that most of the bodily descriptions in the novel are described in the third person (even though it is omnisciently narrated) and that the interpretations laid out in this manuscript are through the lens of clinically documented distortions of the self. (i.e., first-person).

In a rare instance of the first-person description, one of the main characters, Hal Incandenza stands during an interview to explain his low academic achievement and soon realizes that his perceptions of his movements are quite different than those of his audience:

> Please don't think I don't care. "I lookout. Directed my way is horror. I rise from the chair. I see jowls sagging, eyebrows high on trembling foreheads, cheeks bright-white. The chair recedes below me.
> "Sweet mother of Christ," the Director says.
> "I'm fine," I tell them, standing. From the yellow Dean's expression, there's a brutal wind blowing from my direction. Academics' face has gone instantly old. Eight eyes have become blank discs that stare at whatever they see.
> "Good God," whispers Athletics.
> "Please don't worry," I say. "I can explain." I soothe the air with a casual hand.
> Both my arms are pinioned from behind by the Director of Comp., who wrestles me roughly down, on me with all his weight. I taste floor.[…]And his arms." You didn't see it, Tavis. His arms were—"flailing. This sort of awful reaching wriggle. Waggling."
> The group looking briefly at someone outside my sight. (Wallace, 2006, p. 12)

Though Hal attempts to communicate that "nothing is wrong," indeed, he "lies there catatonic, staring" (p. 13).

From a neuroscientific perspective, these various disturbances can be likened to disconnections between body and mind that can rise to disturbances of "self-awareness" in which self-referential processing (i.e., processing information related to the self) is impaired, and indicates some form of disconnection between brain (i.e., what he personally thinks or feels) and behavior. Agency, the experience of being in control of an action or behavior, is a key neuroscientific concept in the novel. Agency is shaped by two significant aspects, fact and sense. Fact alludes to the component of agency in which one purposefully makes things happen by dint of neuroanatomy of voluntary action (i.e., I want to perform an action or accomplish a specific outcome within actions). On the other hand, the sense is "that I am the one who is causing or generating an action. For example, the sense that I am the one who is causing something to move, or that I am the one who is generating a certain thought in my stream of consciousness" (Gallagher, 2000, p. 15). In other words, it implies that the observed outcomes are a direct consequence of initiating and controlling our actions to influence the outside world and the notion of being aware of the outcome (i.e., a conscious experience). Sense of agency is intimately bound with notions of freedom and responsibility traits that give reason to why, if we feel the experience, we consider ourselves initiators of the acts (Haggard and Tsakiris, 2009; Nichols, 2011).

Farrer and Firth, among others, identify that "[a]gency has been assigned a key role in self-consciousness" (Farrer and Frith, 2002, p. 596), and that is why according to Gallagher (2000), self-consciousness is constituted in part by a "minimal self." This direct self-awareness includes self-ownership or the sense that "it is my body that is moving" and self-agency, or the sense that "I am the initiator of the action and thus that I am causally involved in production of that action" (p. 16). In the normal experience of voluntary or willed action, the sense of agency and the sense of ownership coincide and are indistinguishable. In the absence of neural injury, somatosensory signals (e.g., where is my limb in space, what position is it in) align with the neural commands that generate movement (Blakemore et al., 1998). This translates to a feeling of "doing" connected to the brain activity that commands the body to move. If the outcome is predicted accurately, then the sense of agency is stronger (Wegner and Wheatley, 1999). This attenuation permits a distinction between active and passive movements and therefore attributes a sense of agency to actions of individuals. Frith (2005) asserts in "The Neural Basis of Hallucinations and Delusions" that one of the more extensively accepted explanatory models of these symptoms indicates a dysfunction in the "forward model" system (Frith, 1992), whose role consists in predicting the sensory consequences of actions. Due to damage to this comparator of predictions and consequences, individuals cannot discern whether they initiated action and thus conclude that it must be out of the control of an individual.

As fictional characters, Hal and others do not display classic or complete clinical disorders but rather features that represent aspects of these fascinating syndromes. These links can be found with regard to clinical syndromes which include distortions of

agency, self-awareness, and body ownership. Disturbances in an agency are characteristic of certain psychiatric and neurological disorders. For example, deficits of sense of agency are common in patients with schizophrenia (Gallagher, 2005; Jeannerod, 2009). Somatic passivity phenomena (or delusions of alien control) (Oyebode and Davison, 1989) are a hallmark symptom of schizophrenia. The passivity phenomena in schizophrenia are defined as actions, thoughts, and sensations perceived to be under external forces that can be linked to an abnormal sense of agency (Georgieff and Jeannerod, 1998; Frith et al., 2000). Schizophrenia patients with these symptoms report a loss of clear boundaries between the self and others and think/feel that their thoughts and actions are controlled by external forces rather than belonging to them (Blakemore et al., 2002; Lindner et al., 2005; Shergill et al., 2005). Interestingly, for these types of patients, the degree of feeling in "self-touch is as high as the same touch applied by someone else" (Frith, 2005, p. 171). It is not only that they do not perceive it as self-created, but also, that it is intensified by the feeling that it comes from an external entity. In other examples such as that of research of Lindner et al. (2005), schizophrenic patients believed eye movement originated from external forces, a sensation that was linked to a possible "deficit in the perceptual compensation of sensory consequences of one's actions" (p. 1119). This description evokes examples of characters in the novel discussing a dubious sense of control. Dr. Rusk, the psychologist who no player understands says that "on the level of objects and a protective infantile omnipotence where you experience magical thinking and your thoughts and the behavior of objects' relation to your narcissistic wishes, the counterphobia presents as the delusion of some special agency or control to compensate for some repressed wounded inner trauma having to do with absence of control" (Wallace, 2006, p. 550).

Disconnection between body and mind also characterizes a host of neurologic syndromes in which aspects of self-awareness are impaired. Patients who have developed paralysis of one side of their body (i.e., hemiplegia) are often unaware/disconnected from their lost function and deny any motor difficulties as a result. This disorder is commonly referred to as anosognosia, which translates to ἀ- a-, "without," νόσος nosos, "disease" and γνῶσις gnōsis, "knowledge" (Jenkinson et al., 2011; Mograbi and Morris, 2018). One of the most salient delusive phenomena is known as somatoparaphrenia (e.g., patient ascribes ownership of the limb to another person) (Gerstmann, 1942). Following Gerstmann's (1942) definition, for this syndrome to occur, a patient must have (i) acquired contralateral motor deficits, (ii) unawareness of such deficits, and (iii) delusional beliefs regarding the limbs affected by these deficits (Feinberg and Venneri, 2014). These delusional beliefs are specific to body ownership defined by Jenkinson et al. (2018) as the "sense, feeling or judgement that the body belongs to me and is ever present," while other delusions such as *asomatognosia* (e.g., patient behaves as if their limb does not exist) typically entail delusions of *existence, visual self-reflection, and sense of belonging* of the contralateral limb (Jenkinson et al., 2018). Other disturbances can be observed as unilateral spatial neglect following brain injury (Vallar, 1998). Neglect specific to own body of an individual, also known as personal neglect (e.g., Pizzamiglio et al., 1989; Guariglia and Antonucci, 1992; McIntosh et al., 2000; Baas et al., 2011), has been proposed to reflect predominantly right-sided neuroanatomic compromise, including regions such as the postcentral and supramarginal gyri of the parietal cortex (Committeri et al., 2007, 2018). In this disorder, the patient is unable to direct their attention to one side of their body. The extent of this deficit is manifested in everyday activities, such as dressing or grooming, in which patients neglect to dress, brush their hair, or apply make-up on the contralesional side of their bodies (Caggiano and Jehkonen, 2018). Interestingly, anosognosia, asomatognosia, and neglect commonly co-occur and neglect has previously been proposed as an underlying mechanism for anosognosia (Cutting, 1978; Starkstein et al., 1992; Jehkonen et al., 2006). However, anosognosia and accompanying delusions are considered distinct from personal neglect. Indeed, dissociations of these disorders have been reported, and while patients with anosognosia can have delusions of ownership of the contralesional part of their body (e.g., somatoparaphrenia), neglect patients will acknowledge ownership if their attention is properly directed to their contralesional limbs (Bisiach et al., 1986; House and Hodges, 1988; Small and Ellis, 1996; Berti et al., 2005; Jenkinson et al., 2018). These clinical syndromes of body self-awareness may have inspired Wallace in the process of writing *Infinite Jest*. Particularly, one may see neglect as epitomized through the Entertainment film. Viewing the film ceases any corporal attention leading to a possible vegetative state or even death.

Disturbances in the mind–body connection in the novel invite comparison to neurologic disorders in which expression of affect through motor control, particularly through facial expressions, become impaired. Hal, for instance, has lost the ability to express through his body any feelings or sensations. The narration describes that "[f]or a brief moment that Hal will later regard as completely and uncomfortably bizarre, Hal feels at his own face to see whether he is wincing" (Wallace, 2006, p. 342). These anomalies of expression can be observed across multiple clinical disturbances. One such disorder is Parkinson's disease (PD) and atypical Parkinsonian disorders such as progressive supranuclear palsy (PSP). PD, a neurodegenerative disorder characterized by neuronal loss in the substantia nigra (Olanow et al., 2009), manifests in motor disturbances such as bradykinesia, rigidity, tremor, and postural instability (Gelb et al., 1999), while main symptoms of PSP are ocular motor dysfunction, postural instability, akinesia, and cognitive dysfunction (Höglinger et al., 2017). Both PD and PSP patients often show a loss of spontaneous and voluntary facial expressions, commonly referred to as *masked faces* (Buck and Duffy, 1980; Katsikitis and Pilowsky, 1988, 1991; Borod et al., 1990; Saku and Ellgring, 1992; Brozgold et al., 1998; Bowers et al., 2006; Marsili et al., 2014). These deficits can lead to incongruent and reduced/absent facial expressions such as the ones described by Wallace in *Infinite Jest*. Indeed, as opposed to healthy individuals, patients with PD and PSP experience difficulty in expressing facial expressions. These impairments are hypothesized to underlie the degeneration of brain networks that include structures, such as the amygdala and basal ganglia known to be key in emotion processing and motor action (e.g., Davidson et al., 2000; Gross and Thompson, 2006; Aron et al., 2014).

EMBODIMENT IN THE NOVEL

Infinite Jest seems to determine/reflect how the environment affects perception and self-awareness, in the sense of "hereness" of characters at a given moment, and feeling/bodily sensations exist for a particular experience. Overall, a wide range of disruptions (in the vast majority related to emotions and bodily sensations/perceptions) of the novel can be interpreted as leading to the idea that to experience correct self-awareness, there is a need to connect these cognitive processes to a more grounded or lived experience of self. That is, as the narrative is fragmented (although connected in the paradigm of bodily disruptions and lack of correct affective framing), it leaves the question open to how characters can have a complete experience in light of the overpowering technological or entertainment industry. Experiences need to be studied from the disconnection of the organism from its environment. In this case, Wallace reflects on the loneliness of characters that have lost the sense of belonging or communicating with their environment (not only the surroundings but also others). The body in the novel lacks situatedness and the correct arousal of the affective component (as in a dialogue with its surroundings).

The novel is about the sadness, loneliness, or pain of not being connected. As previously discussed, there is a sensation of disconnection with the surroundings that are essential for a complete sense of self-awareness of oneself. The following section is particularly relevant for its intimate discussion on core sections of awareness:

> Hal, who's empty but not dumb, theorizes privately that what passes for hip cynical transcendence of sentiment is really some kind of fear of being really human since to be really human (at least as he conceptualizes it) is probably to be unavoidably sentimental and naive and goo-prone and generally pathetic, is to be in some basic interior way forever infantile, some sort of not-quite-right-looking infant dragging itself analytically around the map, with big wet eyes and froggy-soft skin, huge skull, gooey drool." (Wallace, 2006, p. 694)

The narrative suggests that the disconnection of the "egocentric" or "inner source-point," (Maiese, 2016, p. 10) that is, the "internal self" to bodily sensations ("sentiment and need"), and the disturbances in the situated bodily experience within its environment ("empty mask, anhedonia") are what seems to lie behind self-detachment in the novel. Indeed, lived experiences described in *Infinite Jest* are portrayed as deviated/distorted by disembodied bodies and disturbances perception (connects to theories of the phenomenology of Merleau-Pounty).

The most intimate or personal sections include the influence of embodied cognition or phenomenological description (mainly through tennis and Kate Gompert) demonstrating the importance of the body in experience or for self-awareness: "I was in my body. My body and me were one … my well-armed hand was the secretary of my mind, lithe and responsive and *senza errori*, because I knew myself as a body and was fully inside my little child's body out there" (Wallace, 2006, p. 165).

CONCLUSION

This study has offered a constructive analysis to demonstrate how literature is a well-established vehicle to interpret or illustrate clinical studies/knowledge as well as current social and technological developments of a given time. Societal fascination over the self is also seen through the exploration and discussion of status of the individual in connection to these developments throughout history. Social issues then seem to align with new concepts and discoveries in scientific, technological, and medical fields. One way of interdisciplinary inquiry that reflects upon social curiosity or concerns is through literature as we have attempted to show in this study.

First of all, it has discussed renowned literary authors who have included medical knowledge and discussion in their works throughout history. Second, it has shown parallelism between the fictionalized world in *Infinite Jest* and some major aspects regarding current neuropsychological research on the role of body awareness in shaping the self. In the world of the novel, the state of the body is depicted as malleable, fractioned, deformed, and disconnected from the mind, which gives its characters an unstable sense of self. This is also translated into a dubious sense/feeling of embodiment that has a major role in representing and attributing body and actions of an individual to the self of an individual. By contrasting the representation of bodies and sense of agency in the novel with disorders, such as anosognosia, anorexia, schizophrenia, disembodiment, and lack of sense of agency, we have demonstrated how literature can also supply a resource for examples of body and self not only in a literary way but also reflecting a medical one. These examples also parallel current social issues on the overuse of entertainment and technology, addiction, and the dependency on these external stimuli for identity and self. In this manner, the discussion of these concerns seems to become intertwined with current neuroscientific discoveries observed in this literary text. Knowledge of Wallace on psychological theories detected in his novel has allowed us to explore and connect body self-awareness and mind issues of the period. From a literary standpoint, the adaptation of disorders in body self-awareness and the self also enable Wallace to expose how entertainment and the contemporary way of living has an impact on body self-awareness and the self, thus demonstrating an illustrative relationship between the social, literary, and medical fields of a historical period.

The novel develops a vision of a society whose pursuit of pleasure was shutting itself off from true feeling and experience. A society that was fragmented and not unified/connected but rather constituted by disconnected/self-detached and solipsistic individuals that only share the common ground of alienation and addiction to entertainment. Wallace sets a general medicalized framework (inserting many references to drugs and mental disorders and body disturbances), addiction, and the entertainment industry to explore human emotions and experience. Rather than providing answers or a general diagnosis and "recovery," the dystopian plot opens up to a

more embodied ground for self-awareness. As observed in the literary work, "your self is not only your head but body" (Wallace, 2006, p. 272); Self-awareness then is connected to lived time, the living body and its connection and responses to the environment which emphasizes subjective and lived experience observed in embodied cognition models and in the novel, Wallace shows dislocation/alienation from the globalized but calls for the sincerity of experience. The novel does not "just diagnose a malaise. It proposed a treatment" (Max, 2012, p. 214) from lack of awareness moving away from first-wave cognitive studies of diagnosis as neurological reactions and thus getting closer to understanding through "fiction … what it is to be a fucking *human being*" (McCaffery, 1993, p. 131).

AUTHOR CONTRIBUTIONS

AC took the lead in writing the manuscript. SCh helped shape analysis and provided critical feedback. SCo supervised and provided feedback on the final versions of the manuscript. All authors contributed to the article and approved the submitted version.

ACKNOWLEDGMENTS

This manuscript partially reflects thesis work by Chapman's (2016) Failed entertainment: (Dis)abled bodies, sense of agency and perceptual binding in Infinite Jest (dissertation's thesis). [Seville (ES)]: University of Seville.

REFERENCES

Aron, A. R., Robbins, T. W., and Poldrack, R. A. (2014). Inhibition and the right inferior frontal cortex: one decade on. *Trends Cogn. Sci.* 18, 177–185. doi: 10.1016/j.tics.2013.12.003

Baas, M., De Dreu, C. K., and Nijstad, B. A. (2011). When prevention promotes creativity: the role of mood, regulatory focus, and regulatory closure. *J. Person. Soc. Psychol.* 100, 794–809. doi: 10.1037/a0022981

Berti, A., Bottini, G., Gandola, M., Pia, L., Smania, N., Stracciari, A., et al. (2005). Shared cortical anatomy for motor awareness and motor control. *Science* 309, 488–491. doi: 10.1126/science.1110625

Bisiach, E., Vallar, G., Perani, D., Papagno, C., and Berti, A. (1986). Unawareness of disease following lesions of the right hemisphere: anosognosia for hemiplegia and anosognosia for hemianopia. *Neuropsychologia* 24, 471–482. doi: 10.1016/0028-3932(86)90092-8

Blakemore, S., Wolpert, D. M., and Frith, C. D. (1998). Central cancelation of self-produced tickle sensation. *Nat. Neurosci.* 1, 635–640. doi: 10.1038/2870

Blakemore, S. J., Wolpert, D. M., and Frith, C. D. (2002). Abnormalities in the awareness of action. *Trends Cog. Sci.* 6, 237–242. doi: 10.1016/s1364-6613(02)01907-1

Borod, J. C., Welkowitz, J., Alpert, M., Brozgold, A. Z., Martin, C., Peselow, E., et al. (1990). Parameters of emotional processing in neuropsychiatric disorders: conceptual issues and a battery of tests. *J. Commun. Disord.* 23, 247–271. doi: 10.1016/0021-9924(90)90003-h

Bowers, D., Miller, K., Bosch, W., Gokcay, D., Pedraza, O., Springer, U., et al. (2006). Faces of emotion in Parkinsons disease: micro-expressivity and bradykinesia during voluntary facial expressions. *J. Int. Neuropsychol. Soc.* 12, 765–773. doi: 10.1017/S135561770606111X

Brozgold, A. Z., Borod, J. C., Martin, C. C., Pick, L. H., Alpert, M., and Welkowitz, J. (1998). Social functioning and facial emotional expression in neurological and psychiatric disorders. *Appl. Neuropsychol.* 5, 15–23. doi: 10.1207/s15324826an0501_2

Buck, R., and Duffy, R. J. (1980). Nonverbal communication of affect in brain-damaged patients. *Cortex* 16, 351–362. doi: 10.1016/S0010-9452(80)80037-2

Burke, M., and Troscianko, E. T. (2017). *Cognitive Literary Science*. New York, NY: Oxford University Press.

Burn, S. (2012). "A paradigm for the life of consciousness": closing time in "The Pale King". *Stud. Novel.* 44, 371–388. doi: 10.1353/sdn.2012.0044

Burton, R. (2001). *The Anatomy of Melancholy*. New York, NY: The NY Review of Books.

Caggiano, P., and Jehkonen, M. (2018). The 'neglected' personal neglect. *Neuropsychol. Rev.* 28, 417–435. doi: 10.1007/s11065-018-9394-4

Caracciolo, M. (2016). Cognitive literary studies and the status of interpretation: an attempt at conceptual mapping. *New Lit. History* 47, 187–207. doi: 10.1353/nlh.2016.0003

Chapman, A. (2016). *Failed Entertainment: (Dis)abled Bodies, Sense of Agency and Perceptual Binding in Infinite Jest*. Dissertation's thesis. Seville, ES: University of Seville.

Chaucer, G. (2003). *The Canterbury Tales*. Transl. by N. Coghill. London: Penguin Books.

Cioffi, F. L. (2000). An anguish becomes thing: narrative as performance in David Foster Wallace's Infinite Jest. *Narrative* 8, 161–181. Available online at: http://www.jstor.org/stable/20107209

Committeri, G., Piervincenzi, C., and Pizzamiglio, L. (2018). Personal neglect: a comprehensive theoretical and anatomoclinical review. *Neuropsychologia* 32, 269–279. doi: 10.1037/neu0000409

Committeri, G., Pitzalis, S., Galati, G., Patria, F., Pelle, G., Sabatini, U., et al. (2007). Neural bases of personal and extrapersonal neglect in humans. *Brain* 130, 431–441. doi: 10.1093/brain/awl265

Crane, M. T. (2001). *Shakespeare's Brain: Reading with Cognitive Theory*. New Jersey, NJ: Princeton University Press.

Crespi, B., and Dinsdale, N. (2019). Autism and psychosis as diametrical disorders of embodiment. *Evol. Med. Public Health* 1, 121–138. doi: 10.1093/emph/eoz021

Cutting, J. (1978). Study of anosognosia. *J. Neuro. Neurosurg. Psych.* 41, 548–555. doi: 10.1136/jnnp.41.6.548

Davidson, R. J., Putnam, K. M., and Larson, C. L. (2000). Dysfunction in the neural circuitry of emotion regulation: a possible prelude to violence. *Science*. 289, 591–594. doi: 10.1126/science.289.5479.591

Deleuze, G., and Guattari, F. (2004). *A Thousand Plateaus: Capitalism and Schizophrenia*. Transl. by B. Massumi. London: Continuum.

Farrer, C., and Frith, C. D. (2002). Experiencing oneself vs another person as being the cause of an action: the neural correlates of the experience of agency. *Neuroimage* 15, 596–603. doi: 10.1006/nimg.2001.1009

Feinberg, T. E., and Venneri, A. (2014). Somatoparaphrenia: evolving theories and concepts. *Cortex* 61, 74–80. doi: 10.1016/j.cortex.2014.07.004

Fink, G. R., Marshall, J. C., Halligan, P. W., Frith, C. D., Driver, J., Frackowiak, R. S., et al. (1999). The neural consequences of conflict between intention and the senses. *Brain* 122 (Pt 3), 497–512. doi: 10.1093/brain/122.3.497

Fletcher, P. (1816). *The Purple Island: A Poem*. London: Burton and Briggs.

Frith, C. (1992). *The Cognitive Neuropsychology of Schizophrenia*. London: Psychology Press.

Frith, C. (2005). The neural basis of hallucinations and delusions. *Comp. Rend. Biol.* 328, 169–175. doi: 10.1016/j.crvi.2004.10.012

Frith, C., Blakemore, S., and Wolpert, D. M. (2000). Abnormalities in the awareness and control of action. *Philos. Trans. R. S. B: Biol. Sci.* 355, 1771–1788. doi: 10.1098/rstb.2000.0734

Gallagher, S. (2000). Philosophical conceptions of the self: implications for cognitive science. *Trends Cogn. Sci.* 4, 14–21. doi: 10.1016/s1364-6613(99)01417-5

Gallagher, S. (2005). *How the Body Shapes the Mind*. Oxford: Oxford University Press.

Gelb, D. J., Oliver, E., and Gilman, S. (1999). Diagnostic criteria for Parkinson Disease. *Arch. Neurol.* 56, 33–39. doi: 10.1001/archneur.56.1.33

Georgieff, N., and Jeannerod, M. (1998). Beyond consciousness of external reality: a "who" system for consciousness of action and self-consciousness. *Conscious. Cogn.* 7, 465–477. doi: 10.1006/ccog.1998.0367

Gerstmann, J. (1942). Problem of imperception of disease and of impaired body territories with organic lesions. Relation to body scheme and its disorders. *Arch. Neurol. Psych*. 48, 890–913. doi: 10.1001/archneurpsyc.1942.02290120042003

Gross, J. J., and Thompson, R. A. (2006). "Conceptual foundations," in *Handbook of Emotion Regulation*, ed. J. J. Gross (NY: Guilford Press), 5–18.

Guariglia, C., and Antonucci, G. (1992). Personal and extrapersonal space: a case of neglect dissociation. *Neuropsychol*ogia 30, 1001–1009. doi: 10.1016/0028-3932(92)90051-m

Haggard, P., and Tsakiris, M. (2009). The experience of agency. *Curr. Direct. Psychol. Sci.* 18, 242–246. doi: 10.1111/j.1467-8721.2009.01644.x

Haraway, D. J. (1991). *Simians, Cyborgs, and Women: The Reinvention of Nature*. New York, NY: Routledge.

Hart, F. E. (2001). The epistemology of cognitive literary studies. *Philos. Literat.* 25, 314–334. doi: 10.1353/phl.2001.0031

Höglinger, G., Respondek, G., Stamelou, M., Kurz, C., Josephs, K., Lang, A., et al. (2017). Clinical diagnosis of progressive supranuclear palsy: the movement disorder society criteria. *Mov. Disord*. 32, 853–864. doi: 10.1002/mds.26987

Holland, M. K. (2013). "The art's heart's purpose:" braving the narcissistic loop of David Foster wallace's infinite jest," in *Succeeding Postmodernism: Language and Humanism in Contemporary American Literature* (New York, NY: Bloomsbury), 57–89.

Homer. (1987). *The Iliad*. Transl. by E.V. Rieu. London: Penguin Books.

Homer. (2003). *The Odyssey*. Transl. by M. Hammond. London: Penguin Books.

House, A., and Hodges, J. (1988). Persistent denial of handicap after infarction of the right basal ganglia: a case study. *J. Neuro. Neurosurg. Psych*. 51, 112–115. doi: 10.1136/jnnp.51.1.112

Jackson, T. E. (2000). Questioning interdisciplinarity: cognitive science, evolutionary psychology, and literary criticism. *Poet. Today* 21, 319–347. doi: 10.1215/03335372-21-2-319

Jackson, T. E. (2003). "Literary interpretation" and cognitive literary studies. *Poet. Today* 24, 191–205. doi: 10.1215/03335372-24-2-191

Jeannerod, M. (2009). The sense of agency and its disturbances in schizophrenia: a reappraisal. *Exp. Brain Res.* 192, 527–532. doi: 10.1007/s00221-008-1533-3

Jehkonen, M., Laihosalo, M., and Kettunen, J. (2006). Anosognosia after stroke: assessment, occurrence, subtypes and impact on functional outcome reviewed. *Acta Neuro. Scand*. 114, 293–306. doi: 10.1111/j.1600-0404.2006.00723.x

Jenkinson, P., Moro, V., and Fotopoulou, A. (2018). Definition: asomatognosia. *Cortex*. 101, 300–301. doi: 10.1016/j.cortex.2018.02.001

Jenkinson, P., Preston, C., and Ellis, S. (2011). Unawareness after stroke: a review and practical guide to understanding, assessing, and managing anosognosia for hemiplegia. *J. Clin. Exp. Neuropsychol*. 33, 1079–1093. doi: 10.1080/13803395.2011.596822

Johnson, B. (2004). "Volpone," in *Volpone and Other Plays*, ed. M. Jamieson (London: Penguin Books).

Kalfus, K. (2010). *Wish You Were Here*. The New York Times. Available online at: https://www.nytimes.com/2010/05/30/books/review/Kalfus-t.html (accessed on 28 May, 2010).

Katsikitis, M., and Pilowsky, I. (1988). A study of facial expression in Parkinson's disease using a novel microcomputer-based method. *J. Neurol. Neurosurg. Psychiatry* 51, 362–366. doi: 10.1136/jnnp.51.3.362

Katsikitis, M., and Pilowsky, I. (1991). A controlled quantitative study of facial expression in Parkinson's disease and depression. *J. Nerv. Ment. Dis*. 179, 683–688. doi: 10.1097/00005053-199111000-00006

Kryger, M. (2012). Charles Dickens: impact on medicine and society. *JCSM*. 8, 333–338. doi: 10.5664/jcsm.1930

Lindner, A., Thier, P., Kircher, T. T., Haarmeier, T., and Leube, D. T. (2005). Disorders of agency in schizophrenia correlate with an inability to compensate for the sensory consequences of actions. *Curr. Biol.* 15, 1119–1124. doi: 10.1016/j.cub.2005.05.049

Maiese, M. (2016). *Embodied Selves and Divided Minds*. Oxford: Oxford University Press.

Marsili, L., Agostino, R., Bologna, M., Belvisi, D., Palma, A., Fabbrini, G., et al. (2014). Bradykinesia of posed smiling and voluntary movement of the lower face in Parkinson's disease. *Parkinsonism Relat. Disord*. 20, 370–375. doi: 10.1016/j.parkreldis.2014.01.013

Max, D. T. (2012). *Every Love Story is a Ghost Story: A Life of David Foster Wallace*. New York, NY: Viking.

McCaffery, L. (1993). An interview with David Foster Wallace. *Rev. Contemp. Fict*. 13, 127–150.

McIntosh, R. D., Brodie, E. E., Beschin, N., and Robertson, I. H. (2000). Improving the clinical diagnosis of personal neglect: a reformulated comb and razor test. *Cortex* 36, 289–292. doi: 10.1016/s0010-9452(08)70530-6

Mograbi, D. C., and Morris, R. G. (2018). Anosognosia. *Cortex* 103, 385–386. doi: 10.1016/j.cortex.2018.04.001

Moore, J. W., and Fletcher, P. C. (2012). Sense of agency in health and disease: a review of cue integration approaches. *Conscious. Cogn.* 21, 59–68. doi: 10.1016/j.concog.2011.08.010

Morris, D. (2001). Lived time and absolute knowing: habit and addiction from infinite jest to the phenomenology of spirit. *Clio* 30, 375–415.

Nichols, S. (2011). Experimental philosophy and the problem of free will. *Science* 331, 1401–1403. doi: 10.1126/science.1192931

Olanow, C. W., Stern, M. B., and Sethi, K. (2009). The scientific and clinical basis for the treatment of Parkinson disease. *Neurology* 72(Suppl. 4), S1–S136. doi: 10.1212/WNL.0b013e3181a1d44c

Oyebode, F., and Davison, K. (1989). Epileptic schizophrenia: clinical features and outcome. *Acta Neuro. Scand*. 79, 327–331. doi: 10.1111/j.1600-0447.1989.tb10266.x

Pizzamiglio, L., Judica, A., and Razzano, C., Zoccolotti, P. (1989). Toward a comprehensive diagnosis of visual spatial disorders in unilateral brain damaged patients. *Eval. Psicol*. 5, 199–218.

Richardson, A. (2010). *The Neural Sublime: Cognitive Theories and Romantic Texts. The Neural Sublime: Cognitive Theories and Romantic Texts*. Maryland: The Johns Hopkins University Press.

Richardson, A., and Spolsky, E. (eds.) (2017). *The Work of Fiction Cognition, Culture, and Complexity*. New York, NY: Routledge.

Robb, J., and Harris, O. J. T. (eds.) (2013). *The Body in History: Europe from the Paleolithic to the Future*. New York, NY: Cambridge University Press.

Russell, E. (2010). Some assembly required: the embodied politics of Infinite Jest. *Arizona Q. J. Am. Lit. Cult.Theor*. 66, 147–169. doi: 10.1353/arq.2010.0001

Saku, M., and Ellgring, H. (1992). Emotional reactions to odours in Parkinson's disease A clinical application of ethological methods. *J. Ethol*. 10, 47–52. doi: 10.1007/BF02350186

Shakespeare, W. (1980). *Hamlet*. ed. T. J. B. Spencer. Harmondsworth: Penguin Books.

Shelley, M. (2018). *Frankenstein: the 1818 Text*. London: Penguin Books.

Shergill, S. S., Samson, G., Bays, P. M., Frith, C. D., and Wolpert, D. M. (2005). Evidence for sensory prediction deficits in schizophrenia. *Am. J. Psychiatry* 162, 2384–2386. doi: 10.1176/appi.ajp.162.12.2384

Slade, P. D. (1988). Body image in anorexia nervosa. *Psychopathol. Body Image*. 153, 20–22. doi: 10.1192/S0007125000298930

Small, M., and Ellis, S. (1996). Denial of hemiplegia: an investigation into the theories of causation. *Eur. Neurol.* 36,353–363. doi: 10.1159/000117293

Spolsky, E. (2003). Cognitive literary historicism: a response to Adler and Gross. *Poetics Today* 24, 161–183. doi: 10.1215/03335372-24-2-161

Starkstein, S. E., Fedoroff, J. P., Price, T. R., Leiguarda, R., and Robinson, R. G. (1992). Anosognosia in patients with cerebrovascular lesions: a study of causative factors. *Stroke* 23, 1446–1453.

Vallar, G. (1998). Spatial hemineglect in humans. *Trends Cogn. Sci*. 2, 87–97. doi: 10.1016/s1364-6613(98) 01145-0

Wallace, D. F. (2006). *Infinite Jest*. New York, NY: Back Bay Books.
Wegner, D. M., and Wheatley, T. (1999). Apparent mental causation: sources of the experience of will. *Am. Psychol.* 54, 480–492. doi: 10.1037/0003-066X.54.7.480
Wordsworth, W., and Coleridge, S. T. (2013). *Lyrical Ballads: 1798 and 1802*, ed. F. J. Stafford. (Oxford: OUP).
Zunshine, L. (2010). *Introduction to Cognitive Cultural Studies*. Baltimore: Johns Hopkins University Press.
Zunshine, L. (2017). *The Oxford handbook of cognitive literary studies*. Oxford: Oxford University Press.

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

ORIGINAL RESEARCH
published: 26 November 2021
doi: 10.3389/fpsyg.2021.539596

[Re]moving Bodies – A Shared Diminished Reality Installation for Exploring Relational Movement

Julien Laroche[1,2], Loup Vuarnesson[3,4], Alexandra Endaltseva[5,6], Joseph Dumit[7†] and Asaf Bachrach[2]*

[1] Center for Translational Neurophysiology of Speech and Communication, Italian Institute of Technology, Ferrara, Italy, [2] UMR 7023 Structures Formelles du Langage, St Denis, France, [3] École Nationale Supérieure des Arts Décoratifs, Paris, France, [4] Emotic, Nantes, France, [5] UMR 5044, Centre d'Etude et de Recherche Travail, Organisation, Pouvoir (CERTOP), CNRS, Toulouse, France, [6] INSERM U1276, Centre d'Étude des Mouvements Sociaux (CEMS), EHESS, Paris, France, [7] Science and Technology Studies, University of California, Davis, Davis, CA, United States

OPEN ACCESS

Edited by:
Hanne De Jaegher, University of the Basque Country, Spain

Reviewed by:
Luc Nijs, Ghent University, Belgium
Aleksandra Sherman, Occidental College, United States

***Correspondence:**
Julien Laroche
lajulienroche@gmail.com

†ORCID:
Joseph Dumit
orcid.org/0000-0002-9709-4650

Specialty section:
This article was submitted to Theoretical and Philosophical Psychology, a section of the journal Frontiers in Psychology

Received: *01 March 2020*
Accepted: *05 May 2021*
Published: *26 November 2021*

Citation:
Laroche J, Vuarnesson L, Endaltseva A, Dumit J and Bachrach A (2021) [Re]moving Bodies – A Shared Diminished Reality Installation for Exploring Relational Movement. Front. Psychol. 12:539596. doi: 10.3389/fpsyg.2021.539596

In this article we explore an epistemic approach we name dis/embodiment and introduce "Articulations," an interdisciplinary project bringing together Virtual Reality (VR) designers, cognitive scientists, dancers, anthropologists, and human–machine interaction specialists. According to Erin Manning, our sense of self and other emerges from processes of bodying and relational movement (becoming oneself by moving in relation with the world). The aim of the project is to exploit the potential of multi-person VR in order to explore the intersubjective dynamics of relational movement and bodying, and to do so with scientific, artistic and therapeutic purposes in mind.To achieve this bridge, we bring up a novel paradigm we name "Shared Diminished Reality". It consists in using minimalist representation to instantiate users' bodies in the virtual space. Instead of using humanoid avatars or full body skeletons, we reduce the representation of the moving bodies to three spheres whose trajectories reflect the tracking of the head and the two wrists. This "diminished"virtual rendition of the body-in-movement, we call dis/embodiment. It provides a simple but clear experience of one's own responsive movement in relation to the world and other bodies. It also allows for subtle manipulations of bodies' perceptual and cross-perceptual feedback and simplifies the tracking and the analysis of movements.After having introduced the epistemic framework, the basic architecture, and the empirical method informing the installation, we present and discuss, as a proof-of-concept, some data collected in a situated experiment at a science-art event. We investigate motion patterns observed in different experimental conditions (in which participants either could or could not see the representation of their own hands in the virtual space) and their relation with subjective reports collected. We conclude with reflection on further possibilities of our installation in exploring bodying and relational movement.

Keywords: virtual reality, interpersonal coordination, phenomenology, joint action, social interaction, shared experiences

INTRODUCTION

How can we use Virtual Reality (VR) to better study the role of the movement interactions in our experience of self and other? To approach this question, we propose an epistemic framework, "dis/embodiment," where the body is not taken for granted as physical but is actively constituted and experienced through the relational movement between "the self"and "the other." After explaining this approach, we demonstrate the benefit of using VR as a dis/embodiment technique. We then introduce the concept Shared Diminished Reality (SDR) to describe a minimalist multi-user VR design. To demonstrate the potential of SDR to support dis/embodiment research, we describe "Articulations," an interdisciplinary project at the border of philosophy, art installation and empirical cognitive sciences. We present an analysis of the data collected at the event "Moving Humans: with CREATE" at the Tate Modern, a proof-of-concept situated experiment within a public museum event. Finally, we discuss how "Articulations"and its design contributes to the philosophical and scientific questions about relational movement and bodying.

From the Experiential Body to Bodying Together Through Relational Movements

According to Erin Manning (2013), bodies do not pre-exist as objects within a pre-existing environment (or world); bodies as subjects co-become with the world they move in. Philosopher Brian Massumi builds on this: "*movement moves individuation, and in the process makes that ultimate chunk we call our body an event requiring a verb—bodying*" (Manning, 2013, p. xxiii). Hence moving is at the heart of bodying and worlding, and it is through relational movement that the subject-world is constituted. An epistemic shift from the "movement of bodies"to "bodying"enriches our understanding of movement with emotional and ethico-affective dimensions: how emotions, motions, and consciousness feed into each other (Foolen et al., 2012) in the ongoing formation of what we call "our body." It invites the study of body-in-movement as "*the knowing self [which] is never finished, whole, simply there and original* (...) *able to join with another, to see together without claiming to be another*" (Haraway, 1988, p. 586).

This relational perspective on movement relates to Goffman's (1978) research into how selves are constructed through their performative interactions with objects and persons, and to Blumer's (1969) description of how meanings of self and others are made in interaction. We extend Goffman's (1978) and Blumer's (1969) constructivist approaches, inviting an understanding of the body as an emergent relational happening. We draw our epistemic inspiration from Donna Haraway's notion of "apparatus of bodily production"(Haraway, 1991) which points to how things and bodies—material bodies, semiotic devices, and discursive tools—are always in relational becoming. We also build upon the enactive approach that situates the emergence of embodied experiences in a relational sensorimotor loop between an agent and the world she engages with (Varela et al., 1991) together with others (De Jaegher and Di Paolo, 2007; Fuchs and De Jaegher, 2009).

A number of psychologists and philosophers have long argued that interactions with others in particular take an important place in the constitution of the self and its behavior (Merleau-Ponty, 1960; Vygotsky, 1975; Penuel and Wertsch, 1995; Kyselo, 2014), blurring and making malleable the boundaries between the self and other (e.g.,Mead, 1934). The inherent circularity between the movements of self and other has been described as a self-organized, dynamical process by the enactive approach, which has discussed how movements of interacting agents get coordinated within the interaction process itself (De Jaegher and Di Paolo, 2007; Dumas et al., 2014; Laroche et al., 2014). Manning (2009) names such interpersonal (or, as she puts it, trans-subjective) dynamics "relational movement." According to Manning (2009), by moving together, we are bodying each other. In that sense, the experience of our (and others') bodies, or "bodyings," emerges from relational movements, both during development (Stern, 1985) and during the micro-genesis of our instantaneous interactions (Manning, 2013; Sheets-Johnstone, 1999).

This body we experience as relation can potentially be the locus of a variety of disorders related to disembodiment as well as "relational"syndromes (for schizophrenia, see Varlet et al., 2012; Kyselo, 2016; (for autism, see for example De Jaegher, 2013; Koch et al., 2015; Fitzpatrick et al., 2016; Peper et al., 2016). We speculate that if bodying emerges through relational movement then abilities and disabilities, cognitive and body norms and deviations can be approached through the understanding of the qualities and properties of relational movement. In other words, what has been perceived as a "faulty"or "broken"body might be approached as a relational mismatch. The relational perspective in disability studies understands dis/ability as a relationship between an impairment and environment (Moser, 2006; Shakespeare, 2014; Goodley, 2017). While we do not directly address the topic of dis/ability in this paper, this relational perspective has served us as an inspiration for how the inquiry of bodying and relational movement could be enriched and taken further.

Erin Manning's articulation of "bodying"offers an opportunity for alternative epistemic approaches to the bodies that differ or deviate from medical or heteronormative "norms": bodies (and not only human bodies) which are multiple and compose with each other, bodies which escape fixation. We therefore inquire about ways of studying bodies without taking them for granted: how to stick to relationality and movement in the scientific interrogations of the body? "Not taking for granted"in scientific interrogations of (not only) bodies is what we learn from feminist embodiment approaches (e.g., Haraway, 1988) to compose our conceptual frame. This does not mean a plunge into constructivism. Rather, this move is somewhat close to Latour's (2010)"compositionist manifesto"—a refusal to fall into universalism or relativism in knowing the world and a commitment to rethink the modes of being/acting in the world of scientific facts. For Latour (2010, drawing on Whitehead), bifurcation between the objective and subjective, appearances

and reality is a slippery slope in debates about nature and its futures. Meanwhile, compositionism, putting diverse parts together without threatening their diversity, is "*a search for the Common*" with "*caution and precaution*" (ibid, pp. 487–488; also in Endaltseva and Jerak-Zuiderent, 2021). When it comes to bodies, Latour's claim that "*the task of searching for universality but without believing that this universality is already there, waiting to be unveiled and discovered*" (Latour, 2010, p.474) might start from removing bodies as we know them. This is not to neglect the traces of socio-cultural and hierarchical compositions we embody but to take in that "*what is to be composed may, at any point, be decomposed*" (Latour, 2010, p.474). We propose that bodying and disembodiment are conceptual sisters—moving each other into action, composing and decomposing what is known with, about, and from the body which is always in relational movement with people and things at place. Manning (2009) argues that "*we move not to populate space, not to extend it or to embody it, but to create it*" (Manning, 2009, p.12). Thinking together with Manning and Latour we endeavor to operationalize an epistemic and methodological configuration that does not take bodies for granted, harnessing the movement between disembodiment and bodying as part of the apparatus. We name this conceptual turn "dis/embodiment": by suspending the habit we call "our body," we can provide a novel window (from first and third person perspectives) into the diversity of bodying experiences.

In the next section, we show the benefit of using VR as a technique to approach dis/embodiment, then we propose a materialization of such a methodology in a device called "Articulations"using multi-person VR along minimalist design, a technique we call SDR.

Virtual Reality as a Tool for Dis/embodiment

We propose that VR is a vehicle to approach dis/embodiment. By providing highly realistic and immersive stimuli, VR tools can easily fool the cognitive system and therefore alter the perception of the world and the self. This allows for paradigms that are "*valid and highly ecological without compromising experimental control*" (Loomis et al., 1999). Different kinds and degrees of dis/embodied experiences can then be designed and tested. Moreover, VR tracking systems allow for precise measures of the users' kinematics, without requiring an additional motion capture system. Practically speaking, VR setups are now easy to transport and to install in various venues. Therefore, such setups are easily reproducible "in the field"in experiments with identical parameters and sensorial conditions, as well as quickly customizable to fit the actual environmental configuration (dimensions, length of experimentation, number of participants). What follows is a brief review of empirical research on/in VR which we introduce as a demonstration of its potential for dis/embodiment studies.

The question of the body and its appearance in the virtual world is one of the most important topics in the design of VR environments (Murray and Sixsmith, 1999). Indeed, the constitution of our body's experience seems to be as complex and subject to multiple factors in the virtual world as it is in "normal"reality. The "*Proteus effect*" (Yee and Bailenson, 2007) describes how manipulating the features of a person's avatar modulates her behavior and her affect, with effects lasting even after the immersion. In other words, VR can induce new perceptions of one's own emotional and physical capacity. For instance, people behave more confidently with taller avatars (Yee and Bailenson, 2007; Yee et al., 2009) and they report more negative and aggressive thoughts if their avatars are dressed in black or in Ku Klux Klan outfits (Peña et al., 2009). In general, differential embodiment effectively modulates and changes the way we think, feel and move in socially and culturally patterned ways (Banakou et al., 2013; Kilteni et al., 2013; Peck et al., 2013).

The alteration of a body's position in the virtual space can also create a physical discomfort which subjects are typically trying to compensate for, although their "real" body position was actually comfortable in the first place (De la Peña et al., 2010). The modulation of behavior and affect by bodilyrelated experiences also happens with less anthropomorphic avatars. For instance, robot-like avatars tend to produce a certain feeling of security when facing a dangerous situation (Lugrin et al., 2016). The same goes for the shape taken by the virtual other: the mere choice of avatar skin color has an effect on the willingness of a group of individuals to collaborate (Wallace and Maryott, 2009). In contrast, a balanced combination between visual similarity among members and self-identification enhances social attraction between them, as well as their motivation to contribute to a group task and the task performance itself (Van Der Land et al., 2015). Importantly for our purpose here the feeling of social presence isn't correlated to the anthropomorphic realism of avatars (Nowak and Biocca, 2003).

In sum, with its capacity to modify, transform or compose with the appearance of "real world"percepts (related to ourselves, others, and the environment), VR lets us reconstruct and discover new elements of our self as we perform our identity through self-exploration and role-playing (Turkle, 1995; Taylor, 2006). It allows for experiencing and experimenting with the links between bodying and relational movement in a controlled and replicable manner.

Shared Diminished Reality: Re-moving Bodies

To investigate relational movement and bodying through dis/embodiment, we have developed a multi-person VR installation where the immersed participants are co-present via minimalist dis/embodiment. We call this minimalist paradigm SDR, as we retain basic elements of interactive behaviors yet limit the range of possible actions and the amount of perceivable information. This allows participants to focus their attention on the interaction process and permits us to capture and to analyze body motion more easily and readily.

In order to design a VR environment that challenges what is taken for granted about the body, we chose to stay away from anthropomorphic avatars and to invite users to experience bodying without most of its usual properties. Indeed, extensive research has demonstrated the ease with which we

can still perceive human biological motion (Johansson, 1973), animateness (Heider and Simmel, 1944), emotion (Dittrich et al., 1996), and interactional behaviors (Froese et al., 2014a) from a very reduced set of information consisting of a few moving points. Following a similar logic, we argue for the simplification of the virtual body's appearance to a very minimal form (three spheres in our case). Since both our dis/embodiment approach and our interest in relational movement eschews anthropomorphic realism, we avoid the conundrum of the "neutral." In other words, we do not need to create de-sexualized/de-racialized avatars or make arbitrary or normative choices concerning body shape and appearance. In order to avoid confounding or distracting factors (e.g., independent movement, additional objects, attention-grabbing details), and facilitate real time data streaming, data recording and modeling of interactional behaviors, we also chose to reduce to a minimum the virtual environment itself.

Engaging the main theme of this special issue, SDR allows for experimentally controlled dis/embodiment. For example, we can manipulate the appearance, position, and dynamics of the points representing the avatar, and we can play with the amount of information provided by the environment and the interacting avatars. We can thus modulate and favor the participants' behavior, affect and interaction dynamics that are of direct interest to us. This way, SDR enables us to investigate the possibility for human beings to interact, being deprived of most of the usual communication means such as voice, body posture, facial expression, hand gestures. The underlying assumption here is that getting rid of these sociocultural and linguistic modes of communication would help to highlight (for the experimenters and participants) the dynamics of relational movement that constitute our collective bodying.

What follows in the next part is a detailed presentation of our project "Articulations," which implements a specific instantiation of SDR for the study of bodying and dis/embodiment. We first introduce the project and the techno-scientific framework behind the design and the construction of this device. Then, we describe its modus operandi through a presentation of an art-science public experimentation that took place in June 2019 on the occasion of an science-and-art event at the Tate Modern in London.

Articulations : A Design Project and an Installation

Starting in January 2019, the project "Articulations"has brought together VR designers, cognitive scientists, dancers, anthropologists, and human–machine interaction specialists. The aim was to develop a methodological and technical dis/embodiment installation ("Articulations") that (1) augments and modulates users' sensitivity to the relational experiences of dis/embodiment and bodying and (2) permits scientists to experiment on, observe and interrogate the trans-subjective processes that underlie dis/embodiment and bodying. The central idea has been to offer a collaborative experience by immersing two users, using portable headsets, in an SDR environment, dis/embodying each of them as a set of three floating spheres whose animation is derived from the movement of their head and two hands. Moving between different disciplines and perspectives on the body, the "Articulations"multidisciplinary team engaged in a process of participatory and immersive design. As a result, relationality and bodily lived experiences that the installation aims at highlighting were core parts of its design process itself.

The development of the first "Articulations"prototype took 3 months, after which a group of artists, scientists, and philosophers was invited to spend 3 days in a brainstorming residency with the project team. During the residency, each participant explored the SDR prototype with various partners and also observed others exploring it, thereby soliciting insights from multiple perspectives. We spent over 10 h in facilitated meetings sharing our own experiences and discussing the potentials and limits of the installation in order to improve its design. Following the residency, we organized a first public event with local volunteers (recruited through email and social networks announcements). We developed a quantitative and open-ended questionnaire addressing participants' lived experiences of the device itself, of their own body feelings as well as their relational experiences. We also designed a series of questions to be given to participants in post-experience discussions. The experiences shared by participants informed the next cycle of technical modifications and the creation of a general experimental protocol.

In this protocol, both users have been placed in separate physical squares in order to prevent collisions (**Figure 1**) yet participate in the same space in the virtual environment where they can move, walk and interact in real-time. The virtual space itself has been made exceedingly simple: users were placed on an empty plane, under a luminous and wide blue sky (see **Figure 2**).

Concerning the interaction design, we conceptually draw on Manning's proposition that dance practice orients our attention to the movement of bodying: "*There are an infinity of ways of touching on the more-than that is movement-moving. Dance is one example. What dance gives us are techniques for distilling from the weave of total movement a quality that composes a bodying in motion*" (2013, p. 14). To observe general and spontaneous aspects of relational movements under various conditions of dis/embodiment, the protocol thus invites users to explore the environment (which includes their partner) freely, in a non-goal oriented, non-constrained manner. This is what one would do in Collective Dance Improvisation (CDI), where the activity of the dancers is based on the qualities they experience in moving together (Himberg et al., 2018). Conceptually, dancing is the way we learn to body, that is, to experience and interact with a world.

A SITUATED EXPERIMENT AT THE TATE MODERN

In June 2019, we were invited to present our device during the workshop "Moving Humans" at Tate Modern museum (London, United Kingdom) organized in collaboration with CREATE (University of London). The event was free and open to the public. For the purpose of this event, we created a specific

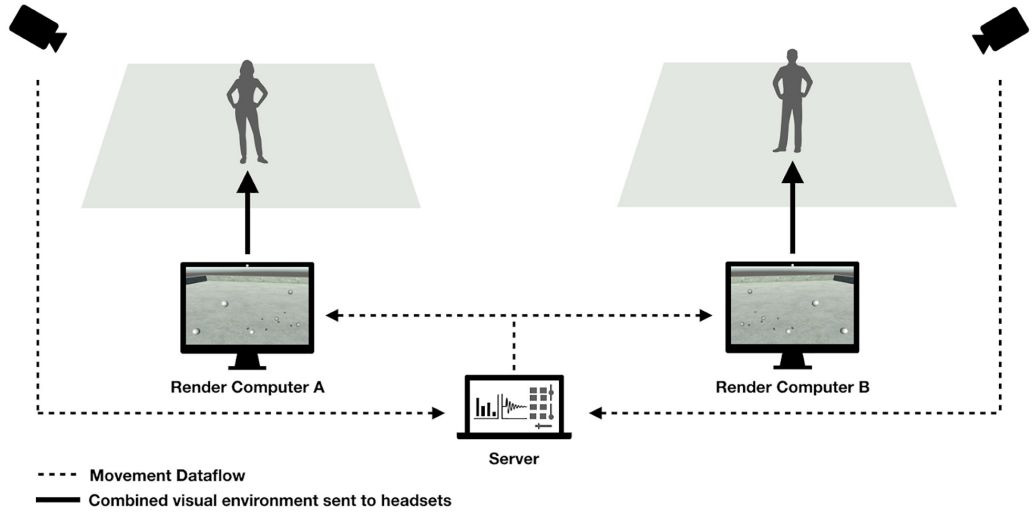

FIGURE 1 | Graphic illustration of the Articulations setup: the two participants are in separate squares but share the same virtual space. HTC Vive lighthouses are installed around the movement space and provide the headsets with localization information. The data is then processed by two dedicated computers, connected via a server program installed on a third computer, that render the virtual scene.

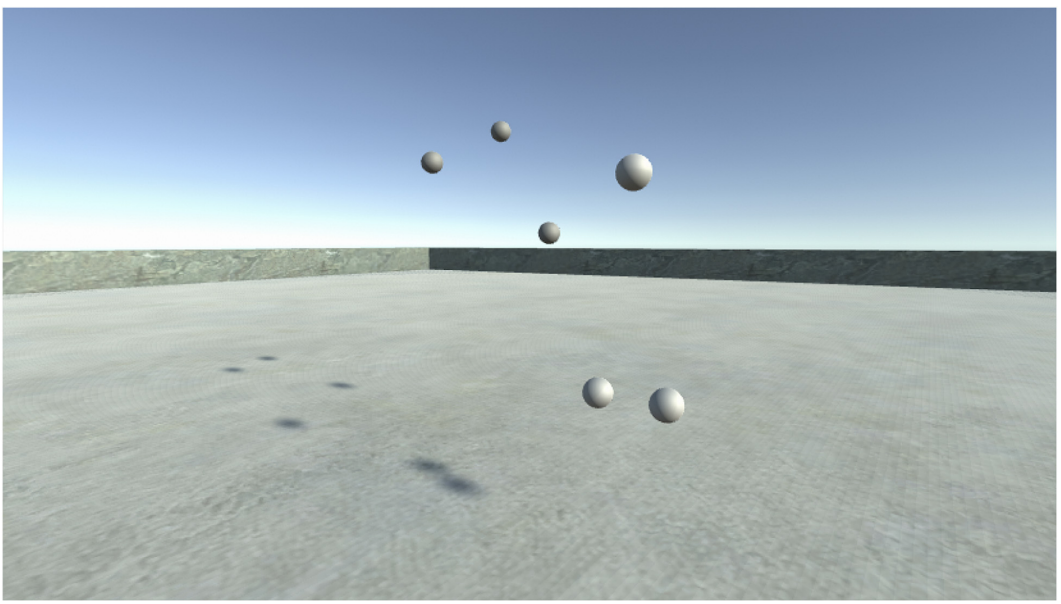

FIGURE 2 | Visualization of the virtual scenery from a third person point of view. We can see the two participants each symbolized as three spheres in the virtual space.

protocol that allowed us to combine an artistic and playful experience with a scientific agenda.

Situated Experimentation

Our interest in introducing a SDR installation into the hybrid art/science event is related to a growing trend of scientific research based on, or targeting, artistic practices (Skov et al., 2018). However, it has a distinctively different logic, inspired by Latour's (2010)"compositionist"search for new ways of acting in the world of scientific facts. We draw upon a specific genre of "acting science"which Latour has developed throughout his collaborative "thought exhibitions"such as Making Things Public: Atmospheres of Democracy or Critical Zones to propose an extension of a "method exhibition"or "data collection exhibition." In other words, we mix data collections, reflection on the data collection, and the demonstration/explication of the data collection for finding a mode of collaborative design *in situ*. Therefore, the protocol which follows does not fit nicely into the category of a lab-experiment, nor is it simply data collection in the field or a scientific observation of an artistic process/event.

The protocol is linked with a particular event where the scientists, artists and public all have their respective "agendas"which might coincide or not (unlike the concept of a subject of a psychology experiment that "serves"the agenda of the experimenter). The protocol integrates multiple contingencies and specificities that cannot be meaningfully replicated, given that it is exactly these singularities that make the event a "happening." For the scientists in the project, the ambition was to be able to collect quantitative data and to be able to observe the effect of slight modification of the experience on the (relational) movement of participants, relating these differences to differences in the lived experience of the same participants. Given the singularity of the event, and since our epistemic approach does want to embrace and not abstract away from real-world complexity, the criteria assumed behind standard statistical/analytical methods are often not met. However, as part of the process of developing custom analysis tools, we chose to apply standard methodology as an attempt to provide connections between this epistemological approach and more canonical experimental results in the literature.

The section below presents the resulting art-science happening prototype, and some ensuing observations. It can be thought of as a report of proof of concept(s) departing from dis/embodiment as an epistemic framework. An epistemic stance of dis/embodiment, we recall, allows a compositionist approach (Latour, 2010) to the body in scientific research (and not only) where the body and its physical composition are not taken for granted. This suggests a possibility to study bodying experiences as they emerge relationally from movement. The event described below works with this suggestion, hypothesizing that a specific form of VR, we called SDR, provides, first, a material framework for dis/embodiment as a concept, and, second, a methodological opportunity to experiment with it. We anticipate that our prototype of SDR, an installation called "Articulations," would allow linking experiential and qualitative investigations of bodying with experimental and quantitative examination of relational movement. Finally, we anticipate a proof of concept on the feasibility of situated art-science experimentation as we present some meaningful data from an installation within the museum, where visitors, fellow exhibitors, and casual passers-by become experiment participants and observers.

Design

A common feature of psychology experiments and (durational) art installations is the use of minimal contrasts (Bateson's "differences that make a difference,"Bateson, 1979). Scientists and artists often introduce subtle changes into an ongoing "event." For the scientist this is a way of disentangling the cognitive correlates of a specific parameter (e.g., perceptual changes). For the artist, minimal differences heighten the aesthetic tension of the work (Massumi, 2015). For the Tate, we chose to produce a 9 min continuous "scenario"within which we introduced, every 3 min, a more or less small change (changing the "condition," in experimental parlance, to allow for contrasts to emerge). We developed two such scenarios, each exploring a slightly different set of questions: the "Transformed Bodies"scenario, which involved changes in avatars' appearance, and the "Mirror"scenario, which involved interacting in presence of a mirror.

In the first scenario ("Transformed Bodies"), the avatars' appearances were slightly different in each condition. In contrast to a baseline situation (where each avatar is represented by spheres corresponding the head and both the hands), a situation was introduced where the spheres representing participants' own hands disappeared (but could still be perceived by the partner); in another one the spheres were displaced, giving the impression of either a stretch or a compression of the arm's length. We expected that eliminating visual feedback of one's own moving body (i.e., of one's own hands) would shift the attention toward the other and the intersubjective space, since the sensorimotor coupling with the other's movements would be the locus of feedback of one's own movements ("perceiving myself through the consequences of my movements on yours"). We hypothesized that this shift would result in enhancing the coordination of movement patterns between the participants. In contrast, we expected that the changes in the perceived size of arms would decrease the ability to coordinate and their associated feelings.

In the second scenario ("Mirror"), we proposed to gauge the changes in behavior occasioned by the perception of a full representation of oneself and the other (all six spheres on the same plane), through a mirror that appeared in the environment. We expected that it would encourage an increase in collaborative movements in order to form shapes together. Furthermore, we wanted to explore the phenomenal and interactional consequences of reducing or adding color cues for self-identification. Specifically, we wondered if an increase in self-other differentiation would lead to an increase or a decrease in experiential and/or behavioral intersubjective coupling. Due to the technical problems, the data of only 6 dyads could be analyzed for this scenario. For this reason, we concentrate this report of proof-of-concept on the "Transformed Bodies"scenario, for which 10 dyads could be analyzed.

Participation

Due to the particularity of the venue, the issue of participation played a singular role in our experiment. In the context of a museum exhibition, introduction of a strict laboratory-like research protocol is not possible. The first challenge is the flow of participants. In a museum, people move and wander freely, governed by their own curiosity. We chose to work with this open ended flow, rather than control it. We invited the audience to stop and watch the running experiments while a screen was showing what was being rendered in the virtual environment. As a consequence, some participants were aware of the idea behind the experiment before doing it, while others were not. Some knew each other and insisted on doing the experiment together while others were partnered randomly. This diversity of cases and contexts had an impact on the collected data. In a certain manner, the participants co-built the experiment with us. We endorse this particularity as part of the logic of the design of "Articulations": the installation is participatory and invites suggestions for co-construction, insight and improvement. We consider the Tate event, therefore, a unique chance to see our

experimental configuration go outside our expectations, making the scientific method itself a part of the experimentation.

In total, 46 participants volunteered for the experiments. 23 dyads were formed based on the order of arrival of participants and were randomly assigned to one of the two experimental scenarios ("Mirror"and "Transformed Bodies"). Due to the technical problems, the data of only 10 dyads (20 participants) could be analyzed in the "Transformed Bodies"(Age: 30.4 ± 11.4), and only for the baseline condition and the condition where the visual feedback of one's own hands was suppressed. The participants of half of the analyzed dyads knew each other well. Only three participants had a very regular dance practice and only one had a regular experience with VR.

Materials

The shared environment was created through three instances of the same custom created Unity program: Two Vive Pro headsets were connected each, via a Vive wireless adapter, to a client-instance running on a dedicated computer rendering the 3D environments at 60 frames per second. The third instance of the program was the server, gathering (and recording) the positional data sent by both headsets and hand trackers and sending them back to the connected clients. The server also acted as a remote controller of the two clients, triggering visual changes and events at specific times, manually or automatically. The VR headset, a wireless Vive Pro, offered a 110° field of view (approximately 90° per eye), with two 1,440 × 1,600 pixels displays. Coupled with two Vive Trackers strapped to the users' wrists, this system allowed us to track and record hand and head movements in a three dimensional space (60 frames per second) as well as the voice during a specific phase of the experience.

Procedure

The two participants were placed in two adjacent (separate but with no occlusion) physical spaces (approximately 4 m square each). Once fitted with the headset, participants discovered an "empty"world with only two visible objects (two spheres) corresponding to the location of the sensors located on their hands (in the article we often refer to these sensors/spheres as "hands"). They were invited to walk around the virtual space and explore its boundaries. After a short while, each participant was able to see three additional mobile spheres corresponding to the hands and head of their partner. The participants were not explicitly told about the mapping of spheres to body parts.

The experimental protocol lasted 9 min and was divided into three"blocks"of 3 min, each corresponding to a different experimental condition and presented in a randomized order, without any interruption between them. In the "Transformed Bodies"experiment, one block served as a baseline, where both avatars were present, each consisting of three spheres of same size and color (**Figure 3**). In another condition (NO-HAND), participants did not see the spheres corresponding to their own hands but each of them could see the spheres representing their partner's hands. In the RESIZE condition, the spheres representing the hands were made either more or less distant from the virtual body than they were in the baseline giving the impression of a stretching or shortening of the virtual "arm."

At the end of the experience, the partner's avatar would disappear, and each participant would find themselves once again alone in the virtual environment, seeing only their own hands. A pre-recorded voice in the headset invited participants to express out loud their experience for about 1 min. The headset microphone allowed us to record their voice while they stayed immersed. After speaking, and once the tracking equipment has been taken off, participants filled in a questionnaire about their lived experience inside the virtual space. As the participants were finished with filling out the questionnaire, they were invited for an open-ended interview with one of the team members. An open-ended interview analysis is beyond the scope of this paper, however. The next section will elaborate on the logic of the questionnaire construction.

Experiential Reports

The goal of the questionnaire was not only to obtain feedback on the device and the experiences it affords but also to check if lived experiences reported by participants could be related to the observable patterns of (individual or relational) bodily behaviors. It was composed of assertions (presented on a tablet) with which participants indicated their agreement on a seven-point Likert scale. We chose this approach (rather than an interrogative form) since it matched our linguistic sources (participants' description of their experiences on earlier occasions) and simplified and homogenized the task for the participants. We were interested in scalar responses as this would allow us to correlate the experiential reports with the movement variables. If we asked different scalar questions (e.g.: "*To what extent did you enjoy the experience?*", "*To what extent did you feel connected to your partner?*"), each question would require an ad-hoc mapping of the specific content of the question onto a scale. The strength of adherence to an assertion, however, can be modeled as a continuous effect which is independent of the content of the assertion and which (from our experience) simplifies the task for participants.

The choice and wording of assertions was grounded in reports collected from the users of the installation during the previous experimentations and collective retreats. Reviewing the answers to the questionnaires and the transcripts of the conversations from these earlier events, we identified statements that identified specific aspects of personal experience that appeared more than once. We extracted representative statements and those prototypical to the categories of the experience most relevant to our research interest (relational movement, bodying, dis/embodiment). Importantly, we formulated the questionnaire by staying as loyal as possible to the wording from the first-person experiences, reformulating them only when the language needed to be specified, clarified, or stylistically adjusted. Adhering to the verbal descriptions provided by the participants themselves was fundamental to keep away from our abstractions and expectations addressing the subjective experiences directly. This approach to building a questionnaire disciplines us to stay "compositionist"(Latour, 2010) with bodying, and not taking bodies as fixed objects.

Most of the questions addressed the general experience (e.g.,"*I would have liked the experience to continue much longer*," "*The*

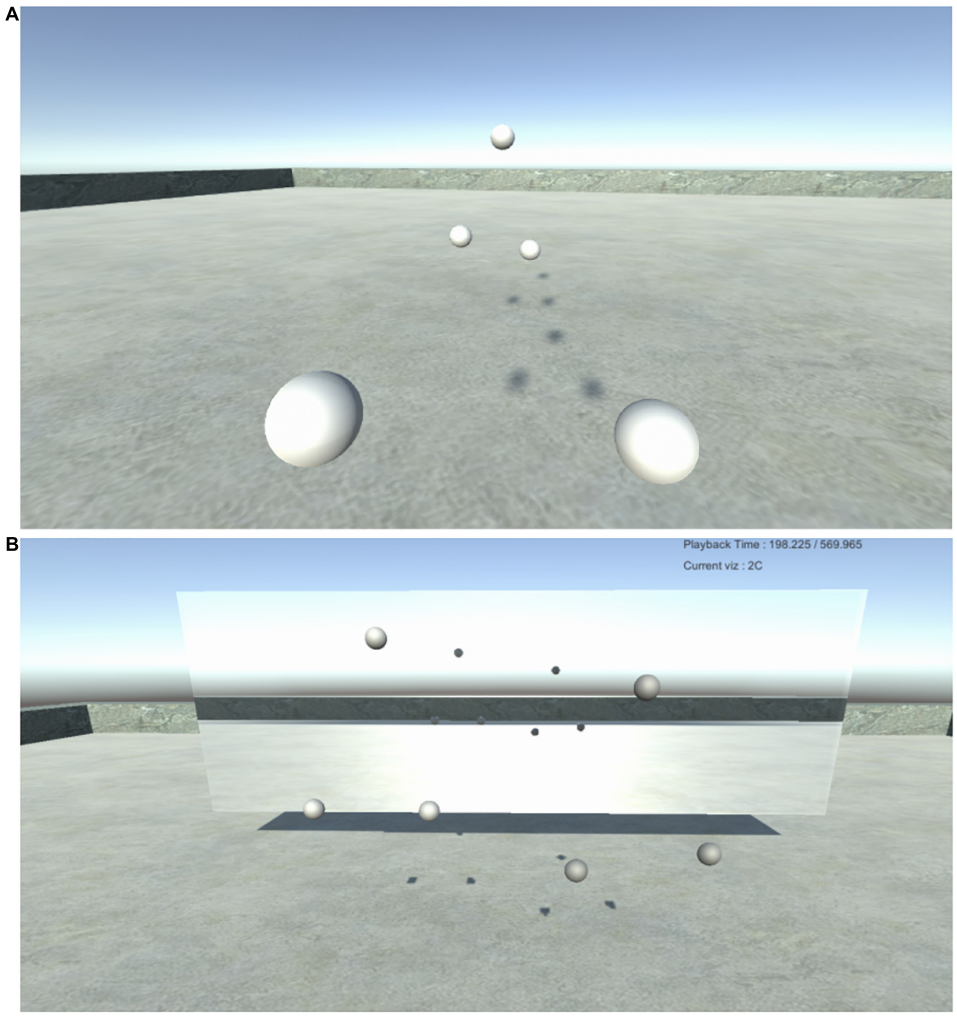

FIGURE 3 | (A) Baseline condition: Visualization of the scenery from a first person point of view. The two closest spheres correspond to the participant's hands. The three spheres at the back of the image correspond to the position of the head and two hands of their partner. Shadows of the floating objects are visible on the ground. **(B)** Mirror scenario: Visualization from a third person point of view. Here both participants are standing side to side in front of their own reflection.

minimality of the setup made me surprisingly more curious and explorative"). A second smaller set of questions targeted the experiences specifically felt during particular experimental conditions. In the analysis presented here we focus on the responses to the latter, in particular to the two questions that inquire about feelings experienced during the analyzed manipulation. Namely the elimination of one's own "hands": (a) "*The absence of my own spheres made me feel more closeness with my partner*"and (b) "*When I could not see my own body I found myself more interacting with my partner*"). We will report how the responses to these two questions related to observed patterns of bodily behavior.

Data Processing, Measurements, and Analysis

In order to quantify how participants moved in the virtual space across conditions, we considered hands and head motion separately. What we mean here by "head"and "hands"are the markers provided by motion capture devices placed on the participants' wrists and the headset. As can be seen in **Figure 4**, vertical head motion happened more sparsely and only intermittently, while participants' hands explored all three dimensions equally. Though we did not collect torso motion data, we repeatedly observed that once immersed, participants mostly did not move the head as an independent effector. Rather, head movement was closely related to the overall motion of the participants trunk in the virtual space. Given these considerations, we took the head's horizontal motion as a proxy for the (horizontal) displacements of the participants in the virtual space (X and Z). For the hands, we took into consideration the overall movement of both hands in the three-dimensional space (X, Y, Z). We first interpolated the original positional time series data at 100 Hz and lowpass filtered them with a second-order Butterworth filter with the cutoff frequency set to 5 Hz in order to eliminate noisy jitter. Then, as a first step

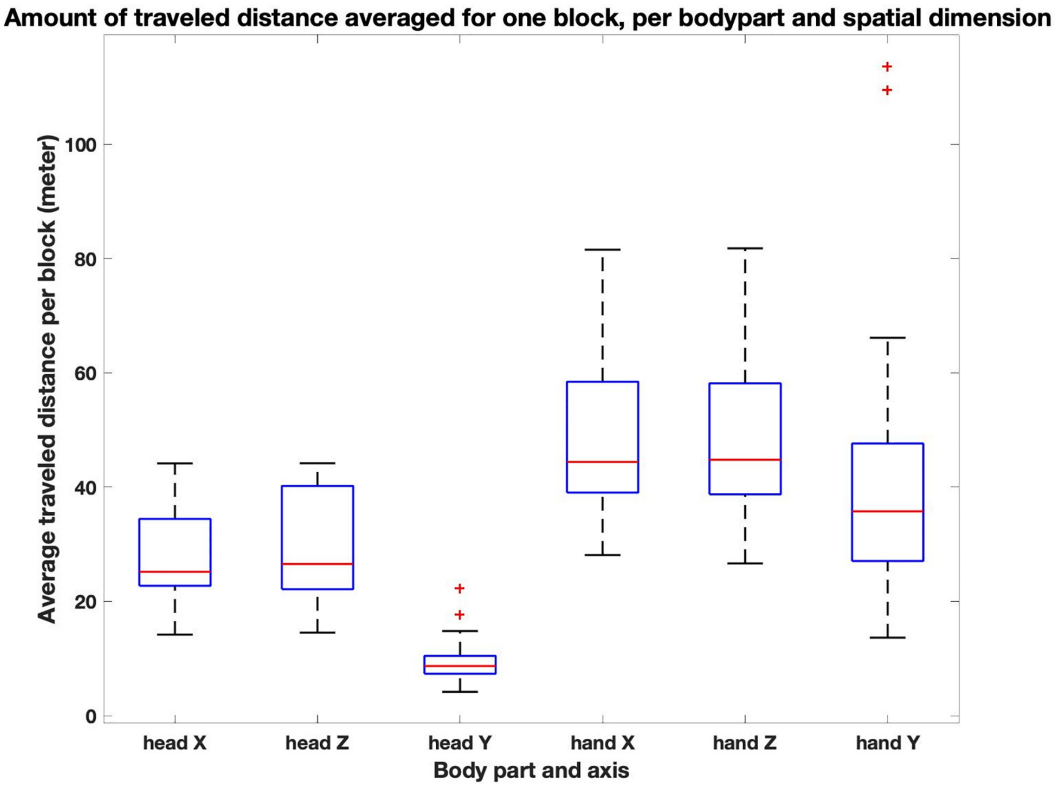

FIGURE 4 | Average amount of distance traveled by block per body part and spatial dimension. In general, distances were quite similar in the X and Z dimension, which reflects the square nature of the virtual space. For the head, displacements along the vertical dimension were significantly smaller than along the horizontal ones (signed-rank Wilcoxon tests; $z = 4.11$, $p < 0.0001$). For the hands, displacements along the vertical dimension are also significantly smaller than along the horizontal ones ($z = 2.74$ and $z = 2.71$, $p < 0.01$ for both vertical vs horizontal comparisons) yet the respective amount of displacements have a comparable range.

in the exploration of the data and with the aim of developing more comprehensive non-linear analysis tools in the future, we computed three types of measurements from the resulting series.

Median Velocity (MV):we first obtained velocity time series by differencing the positional time series of each dimension. Next, each datapoint was squared to obtain unsigned velocity values. Then, we summed the time series of the different dimensions (X and Z for the head, X, Y and Z for both hands), and we computed the square root of each resulting datapoint in order to obtain time series of the overall velocity of the hands and the head, respectively. Finally, we extracted the median of each of the resulting head and hands time series. Thus, for each subject in each condition, and for both the hands and the head, we obtain a singular value reflecting the median velocity of movement.

Relative Head Velocity (RHV): we computed the real-time difference in velocity between partners' head motion in the horizontal plane. To do so, we used the velocity time series described in (1) from both partners for each axis of the horizontal plane (X and Z), to compute a new time series for each axis representing the difference in velocity at every time point between partners. Next, each datapoint in the new time series was squared in order to obtain unsigned difference values. We then summed the time series of the X and Z axis and we computed the square root of each resulting datapoint in order to obtain a time series of the real-time absolute difference between head velocity in the horizontal plane. Finally, we extracted the median of the resulting time series, which provided us with an index of the median instantaneous relative velocity for each pair of participants in each condition.

Motion cross-correlation (MCC): we computed windowed cross-correlations between partner's movements using the overall velocity time series described above and that we used to compute MV. We used moving windows of 2 s (with 1 s overlaps), which can correspond either to a slow movement or to a short sequence of coarticulated gestures. Cross-Correlations were computed over positive lags ranging from 0 to 5 s (by 10 ms overlaps). This allowed us to examine how the movement of each participant related to the present as well as the short-term past of the partner's movement. The 5 s lag value was chosen because movement coordination can take many forms going from real-time synchronization to delayed imitation of a gesture, and can therefore be identifiable at a wide variety of lags (Tschacher et al., 2014; see also **Figure 5**). For this reason, within each window we extracted the peak correlation coefficient whatever was the lag at which it was found. This provided us with a time series indicating for each window the maximum coordination observable between the partner's movement regardless of the delay between these time series. We

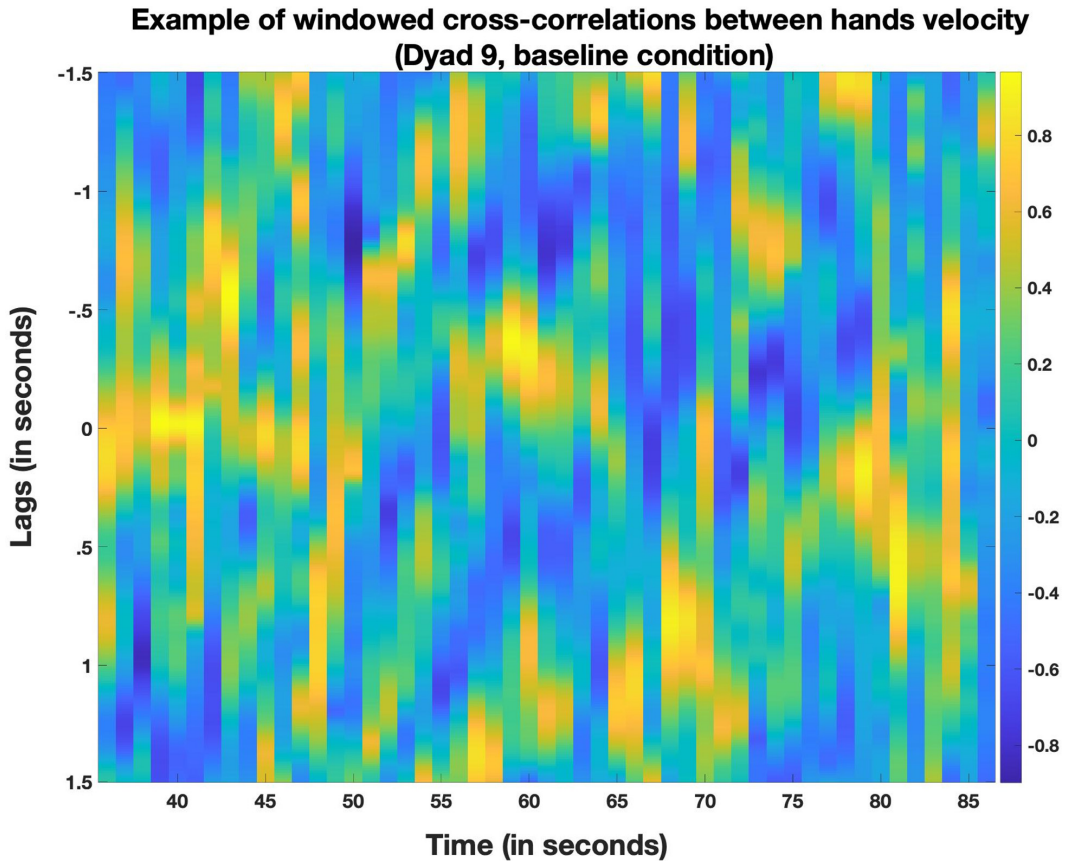

FIGURE 5 | Example of windowed cross-correlations between hands velocity of two partners (dyad 9, baseline condition). Each window is 2 s long. The one second overlap between successive windows entails the creation of a new window at every second of the velocity time series. For every window, a correlation coefficient with the partner time series was computed at every lag between 0 and 5 s, in this example in both time directions (e.g.,correlations with what the partner did up to 5 s earlier as well as later). The resulting correlogram plots the computed coefficient for every window (x axis) at every lag (y axis). The color coding indicates the strength of that correlation coefficient, with cold, most blueish colors indicating strong negative correlation and hot, most yellowish colors indicating strong positive correlations. A positive correlation along the Y-axis 0 indicates synchronous coordination. A positive correlation at negative lags (in the upper part of the plot) indicates that a subject had similar movements to what her partner did earlier (a "follower"). A positive correlation at positive lags (lower part of the plot) indicates that the subject is "leading" (the partner will later make movements that are similar to those the subject makes in the current window). In this example, we can see that both partners first synchronously coordinated (high correlations around lag 0, from 35 to 45 s approximately) before entering a phase that can be described as turn-taking with partners shifting between leading and following.

then used the median of the Z-transformed time series of peak correlation coefficients to provide an index of coordination across time of each participant to his/her partner, in each condition, and for both the head and the hands.

As the distribution of data was not normal, we compared these three indices across conditions (baseline and NO-HAND) with Wilcoxon (signed-rank) tests. For the same reason, relating movement variables to subjective reports, we performed Spearman's rank correlation between the responses to items of the questionnaire that targeted lived experiences during specific experimental conditions and kinematic indices.

Quantitative Observations

Kinematic Measurements

The MCC between partners' hands' velocity was significantly higher in the NO-HAND condition than in the baseline ($z = -2.87$, $p < 0.01$; see **Figure 6**), but no significant difference was found for the head (see **Table 1** for all paired comparisons). In other words, partners coordinated their hands (but not their head) more across time when they could not see their own. There was no significant difference in terms of MV or RHV between the baseline and the NO-HAND condition (see **Table 1**).

Correlation Between Kinematic Measurements and Experiential Questionnaires

We observed a negative correlation between the feeling of closeness induced by the absence of spheres ("*The absence of my own spheres made me feel more closeness with my partner*") and RHV, but this correlation was significant only in the NO-HAND condition ($r = -0.50$, $p < 0.05$; see **Figure 7**). In other words, when their own hands were invisible, participants who moved in space in a more similar velocity, felt closer to each other. The responses to the same question correlated negatively with the participant's head MV, but this correlation was also significant

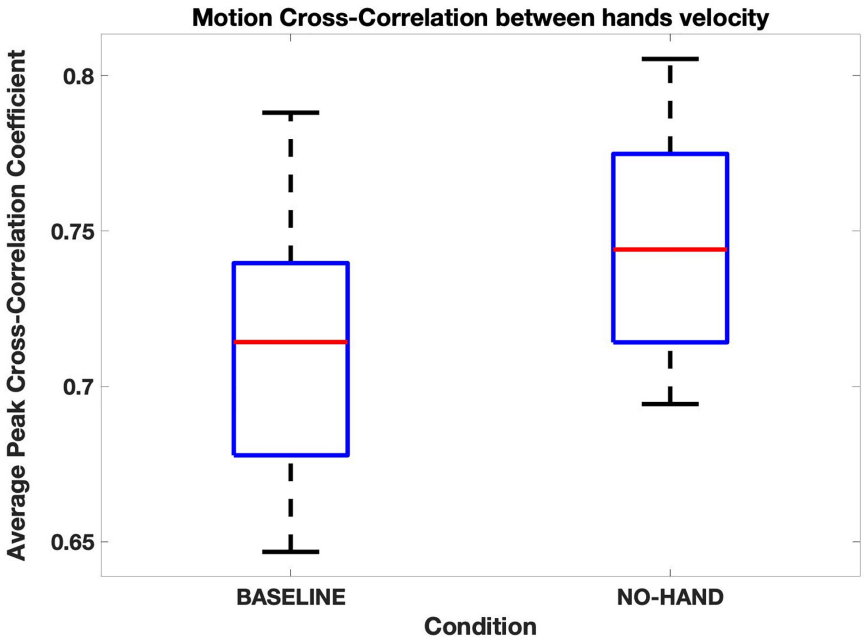

FIGURE 6 | Motion cross-correlations between hands' velocity. Average peak cross-correlation coefficients were significantly higher when participants couldn't see their own hands (NO-HAND) than when they did (BASELINE).

TABLE 1 | Paired Wilcoxon sign-rank tests between BASELINE and NO-HAND conditions for the different kinematics variables.

Variable	Paired Wilcoxon sign-rank test between conditions
HEAD MCC	$z = -1.08$; $p = 0.28$
HANDS MCC	**$z = -2.87$; $p = 0.004$**
HEAD RHV	$z = -0.97$; $p = 0.33$
HEAD MV	$z = -0.30$; $p = 0.77$
HANDS MV	$z = 0.37$; $p = 0.71$

Bold indicates significant comparison.

only in the NO-HAND condition ($r = -0.46$, $p < 0.05$; see **Figure 8**). Importantly, the feeling of closeness in that condition was not significantly connected to the head MV of the partner ($r = -0.25$, $p = 0.28$). These two results indicate that slower self-motion (but not that of the partner) was related to one's feeling of closeness but only when participants could not see their own hands. All other correlation coefficients are reported in **Table 2**.It is important to note that these two movement variables (RHV and head MV) were correlated ($r = 0.65$, $p < 0.01$). This indicates that those who moved faster had more dissimilar head movements velocity than their partner. Given this collinearity, it is not possible to directly parse and gauge the respective and independent contribution of these two variables to the feeling of closeness.

DISCUSSION

In this article we presented the "Articulations"installation for studying relational movement and its situated experimental testing at a science-art event. We first outlined an epistemic framework (dis/embodiment) for the scientific study of the body as an interactive process (bodying). We then introduced the design and operative particularities of SDR as a methodological device for such interrogation, and briefly speculated on its promises for relational approaches to disability. We followed with a description of the "Articulations"SDR installation, where the users are represented most minimalistically, through avatars consisting of three identical mobile spheres corresponding to their head and hands. We finally presented some preliminary results from a situated experimentation at the Tate modern. Our goals at this event were first to test the viability of a SDR set-up to produce and capture (quantitatively as well as qualitatively) bodying through relational movement and dis/embodiment. In addition, we aimed to evaluate the relationship (and our capability to apprehend it) between felt experience and quantitative measures.

We used the specific affordances of SDR to verify if certain manipulations of the visual feedback of one's own movement ("Transformed Bodies") would have significant and differential effects on behaviors, interactions, and lived experiences. We introduced the variations in the virtual environment in pseudo-randomized orders. In the "Transformed Bodies" experiment, the baseline situation was contrasted with one of increased dis/embodiment, where participants could only see the spheres representing their partner. Our expectation was that the disappearance of one's own body (avatar) would draw one's attention to the experience of bodying through relational movement and would favor (measurable) coordination with the partner's movement as their reaction would embody or replace the missing feedback of the participant's own body.

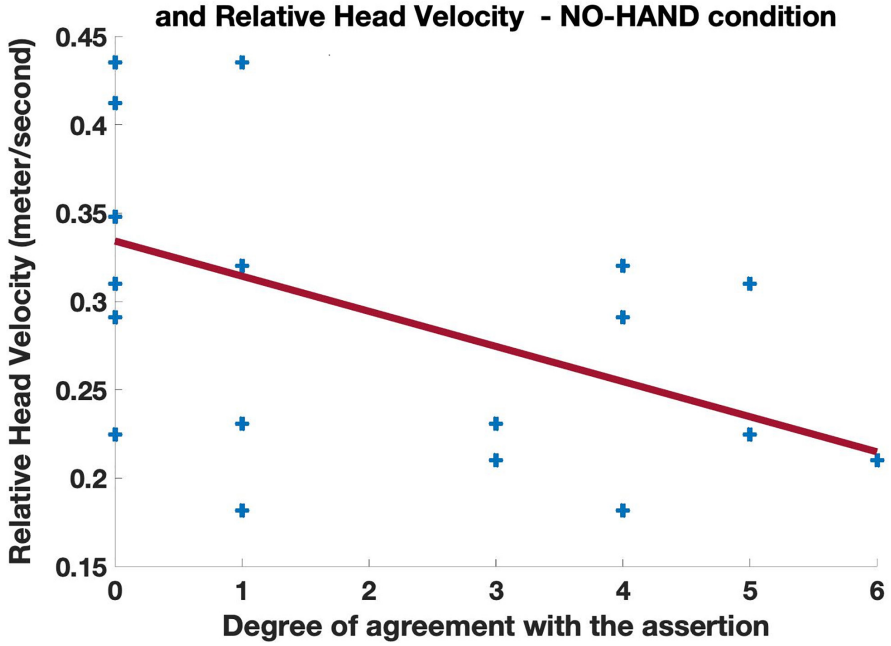

FIGURE 7 | Scatterplot of the relation between the responses to the question "*The absence of my own spheres made me feel more closeness with my partner*" and the Relative Head Velocity. 0 meant total disagreement and 6 total agreement with the assertion, with values in between corresponding to more moderate responses.

Bodying Through Relational Movement

Behavioral results revealed that the velocity of hand movement was more coordinated across partners when the spheres representing participants' own hands disappeared from their view, compared to the baseline situation where they could view their hands. In other words, the disappearance of one's own spheres from one's view, seeing only the other's spheres, seems to have encouraged movement coupling in the dyad. We propose that this specific SDR manipulation provided for a situation where one's own body(ing) was felt through the movement of the other. This dis/embodiment manipulation might have displaced participants' attention toward the other and/or the dynamics of their sensorimotor relationship (in other words, the relationality of their movement). The strengthening of the reciprocal coupling between the participants might have been the consequence of an explicit attempt to communicate and/or a result of unintentional mutual entrainment. We thus interpret the increase in coupling as a reflex of bodying through relational movement. The body here is not a pre-set entity, detached from a pre-existing world; it is constantly composed as a bodying experience where the other becomes the main reference point. The "other"is active in the dis/embodiment experience since the perceptual trace of the consequences of one's own movements in the world can be embodied only through the "other." Thus, the visible body of the "other"co-becomes with the personal pre-reflectively lived body. This is in line with the enactive perspective, which points to the importance of interactional dynamics in the constitution of individual behavior and the coordination of interpersonal behaviors (De Jaegher et al., 2010; Laroche et al., 2014).

Our design and results support those obtained with the cross-perceptual paradigm (Auvray et al., 2009), where participants interact by controlling the most minimalistic agents possible (a unique form of sensory stimulation that signals the presence of the other agent when she is within the receptive field of the participant). This paradigm first demonstrated that individual behaviors can be driven by the collective dynamics that emerge from their interaction (Auvray et al., 2009). Our method proved useful to relate these dynamics to their associated lived experiences (Froese et al., 2014b). Finally, it helped show that even the imitation of bodily behaviors can emerge from interactional dynamics, making the explicit (reflective) perception of one's own bodily expressions unnecessary for imitating each other's bodily behavior (Lenay and Stewart, 2012). Our preliminary results add to this line of inquiry, indicating that the disappearance of one's own body from our perceptual field can strengthen the coupling with the other.

Behavioral Coupling and Lived Experiences

The correlation between behavioral coupling and lived experience (as probed through the questionnaire) suggests that the strengthening of the coupling has an experiential counterpart. Indeed, in the absence of one's own hands, the feeling of closeness was stronger when partners' displacements had more similar velocity (as revealed by RHV). Another finding showed that participants were more likely to signal an enhanced feeling of closeness with their partner in the NO-HAND situation when the velocity of their own displacements (but not those of

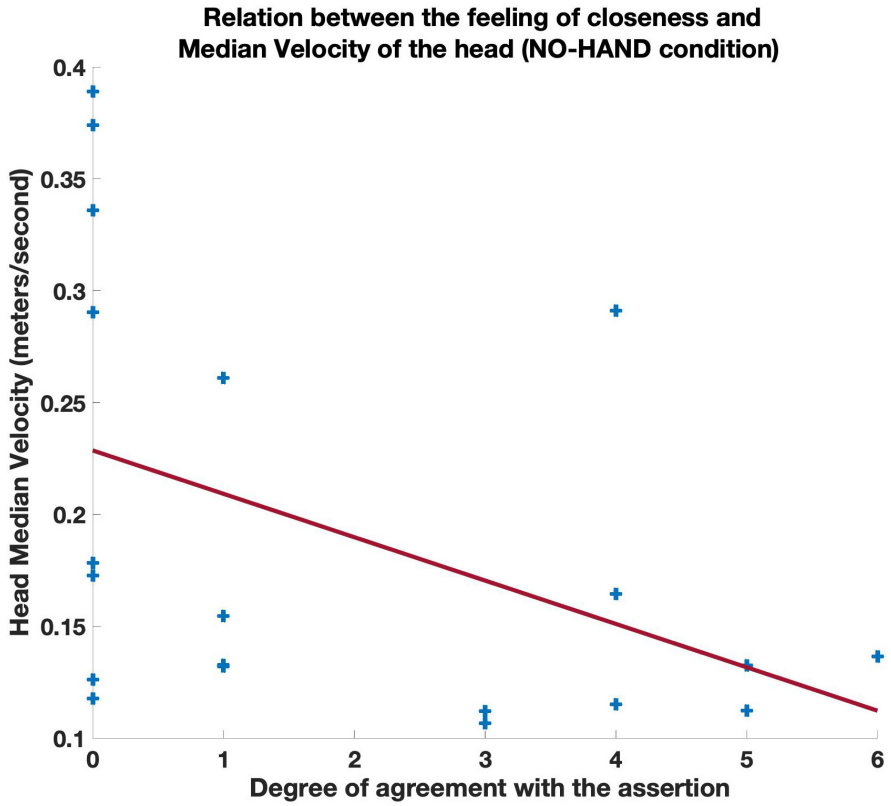

FIGURE 8 | Scatterplot of the relation between the responses to the question "*The absence of my own spheres made me feel more closeness with my partner*" and the Median Velocity of horizontal head displacements. 0 meant total disagreement and 6 total agreement with the assertion, with values in between corresponding to more moderate responses.

the partner) was lower. It thus seems that velocity information is linked to self-other integration in lived experience, whether in relational terms (velocity similarity) or in absolute value (individual velocity).

Velocity information has been shown to convey fundamental information for the recognition of animateness (Tremoulet and Feldman, 2000). It is also determinant in interpersonal movement coordination. Indeed, the mere perception of human motion velocity profiles interferes with the execution of arm movements (Kilner et al., 2007), and the perception of velocity profiles modulates our coordination with an external stimulation (Varlet et al., 2014). Hartmann et al. (2019) showed that hand movement and similarity in the frequency of movements of dancers influenced how observers perceive the degree of interaction between them. Regarding velocity attunement between partners, Słowiński et al. (2016) showed that "*dynamic similarity*" between the individuals' spontaneous motion (according to an index derived from velocity profiles) enhances their interpersonal coordination. It thus seems that the respective velocity profiles of the spheres were important informational cues that led to more or less coordinated patterns through mutual interference and adjustments in the dynamical course of interactions.

Concerning the quality of lived experience associated with interactions, coordination has been recurrently shown to boost affective rapport and/or social connection (Marsh et al., 2009). Our results are notably consistent with those of Sun et al. (2019), who observed positive correlation between head synchrony and feelings of closeness with the partner during conversations in VR. Perceived synchrony has also been shown to increase subjective empathy (Koehne et al., 2016), while objective synchrony between behaviors correlates positively with the subjective estimation of affiliation (Hove and Risen, 2009), rapport (Miles et al., 2009) and closeness/intimacy (Sharon-David et al., 2019). A number of neuroscience studies investigating brain and behavioral activities during interpersonal coordination have guided their analysis with subjective reports (e.g.,Fairhurst et al., 2014). Kokal et al. (2011) showed that participants preferred to be paired with virtual agents that followed them and that this was sufficient to activate brain regions associated with reward processing. Cacioppo et al.'s (2014) results show that synchronized partners seem to be assimilated more, or processed as more similar to the self. Movement synchronization thus seems to enhance shared agency (Pacherie, 2012). Relatedly, movement synchrony has been shown to entail perceived entitativity (Lakens, 2010) and group-bonding (Tarr et al., 2015). Synchrony can thus elicit the feeling of belonging to the same unit. These results resonate with our observation of the correlation between velocity attunement and the feeling of closeness and

TABLE 2 | Spearman rank correlation between experiential reports and kinematic variables.

Variable	Experimental condition	*"The absence of my own spheres made me feel more closeness with my partner"*	*"When I could not see my own body I found myself more interacting with my partner"*
HEAD MCC	BASELINE	r = 0; p = 1	r = −0.17; p = 0.47
	NO-HAND	r = −0.19; p = 0.42	r = 0.14; p = 0.55
HANDS MCC	BASELINE	r = 0.34; p = 0.15	r = 0.11; p = 0.55
	NO-HAND	r = 0.17; p = 0.48	r = 0.11; p = 0.64
HEAD RHV	BASELINE	r = −0.14; p = 0.55	r = −0.11; p = 0.65
	NO-HAND	**r = −0.50; p = 0.02***	r = −0.33; p = 0.16
HEAD MV	BASELINE	r = −0.35; p = 0.13	r = −0.03; p = 0.90
	NO-HAND	**r = −0.46; p = 0.04***	r = −0.35; p = 0.13
HANDS MV	BASELINE	r = −0.28; p = 0.23	r = −0.05; p = 0.84
	NO-HAND	r = −0.10; p = 0.68	r = −0.18; p = 0.46
HEAD MV of the partner	BASELINE	r = 0.23; p = 0.33	r = −0.03; p = 0.90
	NO-HAND	r = −0.25; p = 0.28	r = −0.06; p = 0.81
HANDS MV of the partner	BASELINE	r = −0.04; p = 0.86	r = −0.18; p = 0.44
	NO-HAND	r = −0.08; p = 0.74	r = 0.04; p = 0.87

*Bold and * indicate significant comparison.*

point to a causal link between socio-motor and socio-affective variables. Indeed, the disappearance of one's body increased the coupling with the movement of the other, and it is in this particular condition that we found correlations between synchrony and affiliation. It thus seems that it is the enhancement of interpersonal coordination that entailed stronger feelings of closeness, perhaps by making interacting partners' feelings more similar (or integrated) to each other and therefore feeling more as belonging to the same unit.

Another correlational result links the slowness of a participant's movement to their feeling of closeness to the partner. This result is consistent with Noy et al.'s (2015, Table 7) study. Using the two-dimension mirror game set-up where participants had to imitate each other's improvised hand motion, Noy et al. (2015) found that participants reported moments of togetherness during periods of slower movement though this feeling was not reflected in the objective quantification of kinematic synchrony (co-confident motion). A possible explanation of our result is that slower movements might have allowed participants to stay more attentionally focused on, and therefore aware of and affected by, their partner's movement. Alternatively, a feeling of closeness with the partner could have encouraged the participant to slow down, especially in the challenging NO-HAND condition, to "help"the partner follow or process one's movement, similar to how we introduce disfluencies in conversations to help the listening partner have more "time"to process the information of our speech (Walker et al., 2014). Our result that the feeling of closeness was not correlated with the slowness of the partner argues against an explanation of the feeling of closeness as causally dependent on the partner's slower movement ("*because you move slower I feel more connected to you*").

Here, slowing down as a joint attention mechanism is particularly highlighted when participants don't see their own body since their own kinesthetic "feedback"integrates the movement of the partner. Overall, we favor the hypothesis according to which slower movement might have facilitated movement coordination, which in turn strengthens socio-affective connection (with a positive valence in our experiment and in most studies available in the literature, although it can also lead to stronger connection of an opposite valence; see Di Paolo and De Jaegher, 2012; García and Di Paolo, 2018). Such a hypothesis shall be tested more specifically in future experiments, in particular with the aim of disentangling the pace of movement and the similarity of velocity, which are highly correlated in our preliminary data.

Although correlational, our results expand the literature on the links between motor and socio-affective coordination, that most often refers to rhythmical tasks, to situations where movements are free of any rhythmical constraints. They also argue in favor of the suitability of SDR for the study of relational movement and bodying, and its unique ability to create sensorimotor couplings that can help disentangle the mechanisms of intersubjective coordination. They are also coherent with the enactive perspective where the interpersonal sensorimotor coupling is at the source of experiential attunement, or the mutual incorporation of each other's experience (Fuchs and De Jaegher, 2009). Thinking of our correlational results along with the relational approach in disability studies (Moser, 2006; Goodley, 2017) and Manning's (2009) argument that movement creates space, we speculate that future experiments might build on our findings in searching precise articulation of the link between relational movement and intersubjective experience.

Notes on the Method

The Tate event was a singular happening, and our challenge was to put together a prototype of an epistemological framework that takes into account the rich specificity of the event, with its inherent complexities, while allowing for quantitative observations and qualitative exploration. As a proof of concept,

we gauged the capacity of our SDR set-up to provide methodological opening for a multi-perspective interrogation into the lived experiences of relational movement and bodying through dis/embodiment. We approached interrogation into the lived experiences with a self-report questionnaire which targeted particular experiences. The questionnaire was designed based on the theoretical issues at hand and using as much as possible formulations synthesized from the language used by participants in their feedback during previous sessions. One of the aims of this study was to see if SDR as a methodological device can produce meaningful experiential data.

In this paper we evaluated the link between behavioral measurements and participants' lived experiences. The finding of significant correlations is encouraging since the behavioral measurements we used were rather general (e.g.,overall lateral head or bimanual 3D movement velocity) and did not target specific movement aspects. Plus, the mere fact that the computations were averaged over long (3 min) trials masks the singularity and the variety of the coordination patterns observed in such ill-constrained situations. This prevents us from articulating detailed movement events with particular and locally associated lived experiences. Similarly, the propositions we formulated in the questionnaire addressed lived experiences at the coarse-grain level of specific conditions in their entirety, unlike in-depth elicitation protocols and interviews used by pragmatic phenomenologists that cover seconds (Petitmengin, 2006; De Jaegher et al., 2017; Ollagnier-Beldame and Coupé, 2019). Furthermore, our participants proceeded to the questionnaire only after completing the whole experiment in which the experimental conditions were not clearly isolated in time from one another but rather passed in one continuous run. This contrasts with neurophenomenological approaches where elicitation of experience is more fine-grained and targets trials that are much shorter (e.g.,one perceptual event, see Lutz et al., 2002). In spite of these facts, the statistical significance of our results suggest that our device was successful in revealing bodying experience and creating a link between first-person testimonies and instrumental measurements. One particularly encouraging aspect of the results is that significant correlations were specific to the condition that was targeted by the question. Indeed, the feeling of closeness in absence of the vision of one's own hands was actually significantly (and negatively) correlated with MV and RHV in that very condition, but not with these indices measured in the presence of the said vision of one's own hands. This specificity in the correlations suggests that participants were able to discriminate and refer to the specific experimental conditions post-hoc and retrieve feelings specific to a given condition.

In our ongoing and future research, we intend to attend to the above mentioned problem of granularity, in both the behavioral and experiential domains, by endorsing a more dynamical approach. On the behavioral side, we are currently developing non-linear tools to better capture the temporal evolution of the movement variables such as new methods of recurrence quantification (Lancia and Rosenbaum, 2018). To collect more temporally fine-grained experiential reports, we plan to offer the participants the means to revisit their experiences by navigating a reconstructed video of their interaction. We can then relate specific moments of experience with the local behavioral dynamics. Moreover, this would allow us to interrogate partners about the "same"moments in time. We hope that, with more fine-grained reconstruction of the behaviors and experiences and through combining temporally aligned reports from both partners, we will be able to have deeper insights into the nature of relational movement.

Concerning the diminished aspect, we acknowledge that diminishing perceptual information in VR does not mean that nothing is added in return. The nacreous aspect of the spheres, the marble texture of the ground, and the day-ish light were all designed minimalistically but not in the sense of proposing a mere impoverishment of reality. Rather, we intended to provide a comfortable and pleasant feeling of openness as well as a conducive place for explorative encounters. The spheres also gave inspiration to participants to create patterns and metaphors and filled their imagination to build narratives about their interaction. We saw how many studies explored the impact of body appearance upon how we feel, think, and act. In that case, we can't argue that a body made of spheres is simply a diminution of reality, though it is "minimalist" (Vuarnesson et al., 2021). This "something else" we added will remain as an open question in this paper and should be properly investigated in the future of this project.

We also would like to point out that the way we have implemented SDR in our experiment does not take into account the experiences of bodies with disabilities. We believe that this limitation not only invites us to rethink the logic of participation in the project but also opens a broader question of which knowledge counts in academic research. We recognize that our propositions on the capacities and promises of SDR for the study of bodying have been based on and oriented towards able-bodied experiences. All our participants had a good command of their bodies and did not report motor or cognitive disability experiences. We also recognize that this methodological specification is critically examined in the broader epistemological framework which we draw upon (specifically, the relational model of disability or Haraway's "apparatus of body production"). Though a more detailed development of this point is beyond the scope of this article, we would like to mention here that in our strategy for future research we take better care to include heterogeneous body experiences.

CONCLUSION

In this paper, thinking alongside Erin Manning's concept of bodying, Latour's (2010)"compositionist"approach to knowledge making, and literature on "body production"relational becoming (Haraway, 1988; Moser, 2006), we explored how relational movement underlies the emergence of our bodies and selves. Our working hypothesis is that the experience of our bodily self emerges in our moving relation to people and things which we encounter. We add to this the humble speculation that this process might then provide insights into the phenomena related to dis/embodiment, the experiences of body disorders, and disabilities.

The "Articulations"installation is predicated on the intuition that a task of spontaneous sensorimotor interaction coupled with SDR constitutes a coherent methodological device for this exploration. We have demonstrated that certain manipulations of the virtual space and avatars allow for the modulation of bodily and relational experiences (dis/embodiment) and the observation of the behavioral correlates of this modulation. By testing an installation where two people can freely, spontaneously, and playfully move and interact in the same space, we articulate links between the various forms of dis/embodiment and relational dynamics of movement.

The "Articulations"installation was conceived through a participatory design process, building each feature, from the looks of the virtual movement space to the different conditions to the experience measurement devices, through an iterative exchange among scientists, artists, designers and volunteers who all tested the evolving versions of the installation. This has allowed us to orient the design of the installation towards what we refer to as SDR. SDR constitutes a virtual setup where people can interact in a minimalist environment through avatars whose embodiment is deeply reduced. Such a paradigm allows to retain (or even experientially foreground) the conditions for bodying and relational movements while making kinematics of virtual bodies more easily trackable.

We activated our device at an art-science exhibition where we collected movement data along with experiential reports about experimental manipulations where various forms of dis/embodiment and environmental changes were experienced. Although we have tested our device with only a handful of dyads, which limits the scope of our conclusions, it did offer promising results. As we predicted, suppressing the feedback of the movement of the self increased the coupling with the other. We also observed meaningful links between behavioral measurements and experiential reports that depended on the absence of feedback of one's own hands, at both the individual and the relational level. Without visual feedback of their own movement, members of dyads that had more similar velocities reported a stronger sense of proximity with their partner. This was expected since we created a situation where the visual feedback of one's own movement exists only in the other's movement and in the influence we have on her, which facilitates the orientation of attention toward the coupling process itself. Furthermore, in the same condition participants who moved slower (relatively to other participants) felt closer to their partners. A possible explanation of this last result is that slower movements could enhance the attention toward the partner and the intersubjective relation itself, and/or that the feeling of closeness could bring the participant to slow down as a collaborative strategy. While the causal directionality of this correlation requires further experimentation, it appears that slower movement facilitates the emergence of a relational bodying unit (or "*joint body schema*", Soliman et al., 2015).

The results of this proof of concept show that our SDR methodological device is suitable for exploring the link between bodily lived experiences (especially relational experiences) and movements patterns (whether they refer to patterns of motion observed in individuals or in their relations). Yet, a number of improvements must be considered for further work with our installation: methodological, epistemological, and analytical. Deeper qualitative analysis of the open-ended interviews with the participants and new forms for eliciting experiential reports are needed to enrich our inferences and attend to the specific emotional and social conditions that weave into their experiences; advanced statistical methods are required to improve the detection of relational attunement beyond linear synchronicity or delayed imitation of patterns; semiotic analysis of gestures appropriate for the specific register of spontaneous sensorimotor interactions should be deployed; inviting experiences beyond those of the able-bodied.

Some directions we hope to explore in the future include: qualitative studies of relational experiences with and within SDR; exploration of how dis/embodiment induces novel gesture qualities, and communication strategies; and implications for health and dis/ability fields. More precisely, our interest in the promises of SDR in the context of disability is not about "fixing" or rehabilitating "a body". Rather, we would like to question if relational bodying within the SDR space may inform ethico-affective understanding of dis/ability configurations. We hope that the device we have described here, and its potential for dis/embodiment could contribute to thinking of the ways in which the body is "done"and "undone"collaboratively and to displacing dis/ability from within one's body to the relational ecological network of social and material conditions of its emergence (Moser, 2006; Ginsburg and Rapp, 2013; Shakespeare, 2014; Goodley, 2017).

DATA AVAILABILITY STATEMENT

The datasets generated for this study are available on request to the corresponding author.

ETHICS STATEMENT

Ethical review and approval was not required for the study on human participants in accordance with the local legislation and institutional requirements. Written informed consent to participate in this study was provided by the participants or their legal guardian/next of kin.

AUTHOR CONTRIBUTIONS

LV, AE, and AB collected the results. JL analyzed the data. JL and AB discussed the results. All authors designed the experiment and wrote the manuscript.

FUNDING

This work has been supported by EUR-ArTec and has benefited from French State fundings handled by Agence Nationale de la Recherche for the program "Investissements d'avenir" (ANR-17-EURE-0008). JL was supported by the European Union H2020 – EnTimeMent (FETPROACT-824160) project.

ACKNOWLEDGMENTS

We are deeply thankful to the collaborators without which this research would not have been possible (Dyonisios Zamplaras, François Garnier, Clint Lutes, Thomas Guiffard, Alexandre Coutte, and all those who participated in our early workshops). We would also like to thank Richard Somerville and Ophelia Deroy for inviting us to participate in the Moving Humans event at the Tate Modern. Finally, we are deeply thankful to the reviewers and to the editor for their valuable comments, which helped us improve the text significantly.

REFERENCES

Auvray, M., Lenay, C., and Stewart, J. (2009). Perceptual interactions in a minimalist virtual environment. *New Ideas Psychol.* 27, 32–47. doi: 10.1016/j.newideapsych.2007.12.002

Banakou, D., Groten, R., and Slater, M. (2013). Illusory ownership of a virtual child body causes overestimation of object sizes and implicit attitude changes. *Proc. Natl. Acad. Sci. U. S. A.* 110, 12846–12851. doi: 10.1073/pnas.1306779110

Bateson, G. (1979). in *Mind and Nature: A Necessary Unity*, New York NY: Prentice Hall.

Blumer, H. (1969). *Symbolic Interactionism: Perspective and Method*. California: University of California Press.

Cacioppo, S., Zhou, H., Monteleone, G., Majka, E. A., Quinn, K. A., Ball, A. B., et al. (2014). You are in sync with me: neural correlates of interpersonal synchrony with a partner. *Neuroscience* 277, 842–858. doi: 10.1016/j.neuroscience.2014.07.051

De Jaegher, H. (2013). Embodiment and sense-making in autism. *Front. Integr. Neurosci.* 7:15.

De Jaegher, H., and Di Paolo, E. (2007). Participatory sense-making. *Phenomenol. Cogn. Sci.* 6, 485–507.

De Jaegher, H., Di Paolo, E., and Gallagher, S. (2010). Can social interaction constitute social cognition? *Trends Cogn. Sci.* 14, 441–447. doi: 10.1016/j.tics.2010.06.009

De Jaegher, H., Pieper, B., Clénin, D., and Fuchs, T. (2017). Grasping intersubjectivity: an invitation to embody social interaction research. *Phenomenol. Cogn. Sci.* 16, 491–523. doi: 10.1007/s11097-016-9469-8

De la Peña, N., Weil, P., Llobera, J., Giannopoulos, E., Pomés, A., Spanlang, B., et al. (2010). Immersive journalism: immersive virtual reality for the first-person experience of news. *Presence* 19, 291–301. doi: 10.1162/pres_a_00005

Di Paolo, E. A., and De Jaegher, H. (2012). The Interactive Brain Hypothesis. *Front. Hum. Neurosci.* 6:163.

Dittrich, W. H., Troscianko, T., Lea, S. E., and Morgan, D. (1996). Perception of emotion from dynamic point-light displays represented in dance. *Perception* 25, 727–738. doi: 10.1068/p250727

Dumas, G., Laroche, J., and Lehmann, A. (2014). Your body, my body, our coupling moves our bodies. *Front. Hum. Neurosci.* 8:1004.

Endaltseva, A., and Jerak-Zuiderent, S. (2021). Embodiment work in ethnographic collaborations: Composition, movement, and pausing within the Multiple Sclerosis Society in Russia. *Sci. Technol. Stud.* 2021:5.

Fairhurst, M. T., Janata, P., and Keller, P. E. (2014). Leading the follower: an fMRI investigation of dynamic cooperativity and leader–follower strategies in synchronization with an adaptive virtual partner. *Neuroimage* 84, 688–697. doi: 10.1016/j.neuroimage.2013.09.027

Fitzpatrick, P., Frazier, J. A., Cochran, D. M., Mitchell, T., Coleman, C., and Schmidt, R. C. (2016). Impairments of social motor synchrony evident in autism spectrum disorder. *Front. Psychol.* 7:1323.

Foolen, A., Lüdtke, U. M., Racine, T. P., and Zlatev, J. (eds) (2012). *Moving Ourselves, Moving Others: Motion And Emotion In Intersubjectivity, Consciousness And Language*. Berlin: John Benjamins Publishing Company.

Froese, T., Iizuka, H., and Ikegami, T. (2014a). Embodied social interaction constitutes social cognition in pairs of humans: a minimalist virtual reality experiment. *Scientific Rep.* 4:3672.

Froese, T., Iizuka, H., and Ikegami, T. (2014b). Using minimal human-computer interfaces for studying the interactive development of social awareness. *Front. Psychol.* 5:1061.

Fuchs, T., and De Jaegher, H. (2009). Enactive intersubjectivity: Participatory sense-making and mutual incorporation. *Phenomenol. Cogn. Sci.* 8, 465–486. doi: 10.1007/s11097-009-9136-4

García, E., and Di Paolo, E. A. (2018). Embodied Coordination and Psychotherapeutic Outcome: Beyond Direct Mappings. *Front. Psychol.* 9:1257.

Ginsburg, F., and Rapp, R. (2013). Disability worlds. *Annu. Rev. Anthropol.* 42, 53–68.

Goffman, E. (1978). *The Presentation Of Self In Everyday Life*. Harmondsworth: Doubleday.

Goodley, D. (2017). *Disability Studies: An Interdisciplinary Introduction*, 2nd Edn. Sage: Sage Publications Ltd.

Haraway, D. J. (1988). Situated Knowledges: The Science Question in Feminism and the Privilege of Partial Perspective. *Feminist Stud.* 14, 575–599. doi: 10.2307/3178066

Haraway, D. J. (1991). *Simians, Cyborgs and Women: The Reinvention of Nature*. New York: Routledge.

Hartmann, M., Mavrolampados, A., Allingham, E., Carlson, E., Burger, B., and Toiviainen, P. (2019). Kinematics of perceived dyadic coordination in dance. *Scientific Rep.* 9, 1–14.

Heider, F., and Simmel, M. (1944). An experimental study of apparent behavior. *Am. J. Psychol.* 57, 243–259. doi: 10.2307/1416950

Himberg, T., Laroche, J., Bigé, R., Buchkowski, M., and Bachrach, A. (2018). Coordinated interpersonal behaviour in collective dance improvisation: the aesthetics of kinaesthetic togetherness. *Behav. Sci.* 8:23. doi: 10.3390/bs8020023

Hove, M. J., and Risen, J. L. (2009). It's all in the timing: Interpersonal synchrony increases affiliation. *Soc. Cogn.* 27, 949–960. doi: 10.1521/soco.2009.27.6.949

Johansson, G. (1973). Visual perception of biological motion and a model for its analysis. *Percept. Psychophys.* 14, 201–211. doi: 10.3758/bf03212378

Kilner, J., Hamilton, A. F. D. C., and Blakemore, S. J. (2007). Interference effect of observed human movement on action is due to velocity profile of biological motion. *Soc. Neurosci.* 2, 158–166. doi: 10.1080/17470910701428190

Kilteni, K., Bergstrom, I., and Slater, M. (2013). Drumming in immersive virtual reality: the body shapes the way we play. *IEEE Transact. Visual. Comput. Graph.* 19, 597–605. doi: 10.1109/tvcg.2013.29

Koch, S. C., Mehl, L., Sobanski, E., Sieber, M., and Fuchs, T. (2015). Fixing the mirrors: A feasibility study of the effects of dance movement therapy on young adults with autism spectrum disorder. *Autism* 19, 338–350. doi: 10.1177/1362361314522353

Koehne, S., Hatri, A., Cacioppo, J. T., and Dziobek, I. (2016). Perceived interpersonal synchrony increases empathy: insights from autism spectrum disorder. *Cognition* 146, 8–15. doi: 10.1016/j.cognition.2015.09.007

Kokal, I., Engel, A., Kirschner, S., and Keysers, C. (2011). Synchronized drumming enhances activity in the caudate and facilitates prosocial commitment - If the rhythm comes easily. *PLoS One* 6:e27272. doi: 10.1371/journal.pone.0027272

Kyselo, M. (2014). The body social: An enactive approach to the self. *Front. Psychol.* 5:986.

Kyselo, M. (2016). The enactive approach and disorders of the self-the case of schizophrenia. *Phenomenol. Cogn. Sci.* 15, 591–616. doi: 10.1007/s11097-015-9441-z

Lakens, D. (2010). Movement synchrony and perceived entitativity. *J. Exp. Soc. Psychol.* 46, 701–708. doi: 10.1016/j.jesp.2010.03.015

Lancia, L., and Rosenbaum, B. (2018). Coupling relations underlying the production of speech articulator movements and their invariance to speech rate. *Biol. Cyb.* 112, 253–276. doi: 10.1007/s00422-018-0749-y

Laroche, J., Berardi, A. M., and Brangier, E. (2014). Embodiment of intersubjective time: relational dynamics as attractors in the temporal coordination of interpersonal behaviors and experiences. *Front. Psychol.* 5:1180.

Latour, B. (2010). An attempt at a 'Compositionist Manifesto'. *N. Lit. Hist.* 41, 471–490.

Lenay, C., and Stewart, J. (2012). Minimalist approach to perceptual interactions. *Front. Hum. Neurosci.* 6:98.

Loomis, J. M., Blascovich, J. J., and Beall, A. C. (1999). Immersive virtual environment technology as a basic research tool in psychology. *Behav. Res. Methods Instr. Comput.* 31, 557–564. doi: 10.3758/bf03200735

Lugrin, J. L., Polyschev, I., Roth, D., and Latoschik, M. E. (2016). "Avatar anthropomorphism and acrophobia," in *Proceedings of the 22nd ACM Conference on Virtual Reality Software and Technology*, (New York: Routledge). 315–316.

Lutz, A., Lachaux, J. P., 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.* 99, 1586–1591. doi: 10.1073/pnas.032658199

Manning, E. (2009). *Relationscapes: Movement, art, philosophy*. Delhi: MIT Press.

Manning, E. (2013). *Always More Than One: Individuation's Dance*. Duke: Duke University Press.

Marsh, K. L., Richardson, M. J., and Schmidt, R. C. (2009). Social connection through joint action and interpersonal coordination. *Topics Cogn. Sci.* 1, 320–339. doi: 10.1111/j.1756-8765.2009.01022.x

Massumi, B. (2015). *Politics Of Affect*. Polity Press.

Mead, G. H. (1934). *Mind, Self And Society*. Chicago: University of Chicago Press.

Merleau-Ponty, M. (1960). *Le Philosophe Et Son Ombre*. Chicago: University of Chicago Press.

Miles, L. K., Griffiths, J. L., Richardson, M. J., and Macrae, C. N. (2010). Too late to coordinate: Contextual influences on behavioral synchrony. *Europ. J. Soc. Psychol.* 40, 52–60.

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

Moser, I. (2006). Disability and the promises of technology: Technology, subjectivity and embodiment within an order of the normal. *Inform. Commun. Soc.* 9, 373–395. doi: 10.1080/13691180600751348

Murray, C. D., and Sixsmith, J. (1999). The corporeal body in virtual reality. *Ethos* 27, 315–343. doi: 10.1525/eth.1999.27.3.315

Nowak, K. L., and Biocca, F. (2003). The effect of the agency and anthropomorphism on users' sense of telepresence, copresence, and social presence in virtual environments. *Presence* 12, 481–494. doi: 10.1162/105474603322761289

Noy, L., Levit-Binun, N., and Golland, Y. (2015). Being in the zone: physiological markers of togetherness in joint improvisation. *Front. Hum. Neurosci.* 9:187.

Ollagnier-Beldame, M., and Coupé, C. (2019). Meeting You for the First Time: Descriptive Categories of an Intersubjective Experience. *Construct. Found.* 14:2019.

Pacherie, E. (2012). "The phenomenology of joint action: Self-agency vs. joint-agency," in *Joint attention: New developments*, ed. A. Seemann (Delhi: MIT Press), 343–389.

Peck, T. C., Seinfeld, S., Aglioti, S. M., and Slater, M. (2013). Putting yourself in the skin of a black avatar reduces implicit racial bias. *Conscious. Cogn.* 22, 779–787. doi: 10.1016/j.concog.2013.04.016

Peña, J., Hancock, J. T., and Merola, N. A. (2009). The priming effects of avatars in virtual settings. *Commun. Res.* 36, 838–856. doi: 10.1177/0093650209346802

Penuel, W. R., and Wertsch, J. V. (1995). Vygotsky and identity formation: A sociocultural approach. *Educat. Psychol.* 30, 83–92. doi: 10.1207/s15326985ep3002_5

Peper, C. L. E., van der Wal, S. J., and Begeer, S. (2016). Autism in action: reduced bodily connectedness during social interactions? *Front. Psychol.* 7:1862.

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

Shakespeare, T. (2014). *Disability Rights and Wrongs Revisited*. New York: Routledge.

Sharon-David, H., Mizrahi, M., Rinott, M., Golland, Y., and Birnbaum, G. E. (2019). Being on the same wavelength: Behavioral synchrony between partners and its influence on the experience of intimacy. *J. Soc. Person. Relat.* 36, 2983–3008. doi: 10.1177/0265407518809478

Sheets-Johnstone, M. (1999). *The primacy of movement*. Sage: John Benjamins Pub.

Skov, M., Vartanian, O., Martindale, C., and Berleant, A. (2018). *Neuroaesthetics*. New York: Routledge.

Słowiński, P., Zhai, C., Alderisio, F., Salesse, R., Gueugnon, M., Marin, L., et al. (2016). Dynamic similarity promotes interpersonal coordination in joint action. *J. Roy. Soc. Interf.* 13:20151093. doi: 10.1098/rsif.2015.1093

Soliman, T. M., Ferguson, R., Dexheimer, M. S., and Glenberg, A. M. (2015). Consequences of joint action: entanglement with your partner. *J. Exp. Psychol. Gen.* 144, 873–888. doi: 10.1037/xge0000089

Stern, D. N. (1985). *The Interpersonal World of the Infant*. New York: Routledge.

Sun, Y., Shaikh, O., and Won, A. S. (2019). Nonverbal synchrony in virtual reality. *PLoS One* 14:e0221803. doi: 10.1371/journal.pone.0221803

Tarr, B., Launay, J., Cohen, E., and Dunbar, R. (2015). Synchrony and exertion during dance independently raise pain threshold and encourage social bonding. *Biol. Lett.* 11:20150767. doi: 10.1098/rsbl.2015.0767

Taylor, T. L. (2006). *Play Between Worlds: Exploring Online Game Culture*. Delhi: MIT Press.

Tremoulet, P. D., and Feldman, J. (2000). Perception of animacy from the motion of a single object. *Perception* 29, 943–951. doi: 10.1068/p3101

Tschacher, W., Rees, G. M., and Ramseyer, F. (2014). Nonverbal synchrony and affect in dyadic interactions. *Front. Psychol.* 5:1323.

Turkle, S. (1995). *Life on the Screen: Identity in the Age of the Internet*. Delhi: MIT Press.

Van Der Land, S. F., Schouten, A. P., Feldberg, F., Huysman, M., and vanden Hooff, B. (2015). Does avatar appearance matter? How team visual similarity and member–avatar similarity influence virtual team performance. *Hum. Commun. Res.* 41, 128–153. doi: 10.1111/hcre.12044

Varela, F. J., Rosch, E., and Thompson, E. (1991). *The Embodied Mind*. Cambridge, MA: MIT Press.

Varlet, M., Coey, C. A., Schmidt, R. C., Marin, L., Bardy, B. G., and Richardson, M. J. (2014). Influence of stimulus velocity profile on rhythmic visuomotor coordination. *J. Exp. Psychol.* 40:1849. doi: 10.1037/a0037417

Varlet, M., Marin, L., Raffard, S., Schmidt, R. C., Capdevielle, D., Boulenger, J. P., et al. (2012). Impairments of social motor coordination in schizophrenia. *PLoS One* 7: e29772. doi: 10.1371/journal.pone.0029772

Vuarnesson, L., Zamplaras, D., Laroche, J., Dumit, J., Lutes, C., Bachrach, A., et al. (2021). Shared Diminished Reality: a new framework for the study of embodied intersubjectivity. *Front. Virt. Real.* 2:2021.

Vygotsky, L. S. (1975). *Thought and Language*. Cambridge MA: Cambridge University.

Walker, E. J., Risko, E. F., and Kingstone, A. (2014). Fillers as signals: Evidence from a question–answering paradigm. *Disc. Proc.* 51, 264–286. doi: 10.1080/0163853x.2013.862478

Wallace, P., and Maryott, J. (2009). The impact of avatar self-representation on collaboration in virtual worlds. *Innovate* 5:5.

Yee, N., and Bailenson, J. (2007). The Proteus effect: The effect of transformed self-representation on behavior. *Hum. Commun. Res.* 33, 271–290. doi: 10.1111/j.1468-2958.2007.00299.x

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

Conflict of Interest: LV was employed by company Emotic.

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.

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.

Advantages of publishing in Frontiers

OPEN ACCESS
Articles are free to read for greatest visibility and readership

FAST PUBLICATION
Around 90 days from submission to decision

HIGH QUALITY PEER-REVIEW
Rigorous, collaborative, and constructive peer-review

TRANSPARENT PEER-REVIEW
Editors and reviewers acknowledged by name on published articles

Frontiers
Avenue du Tribunal-Fédéral 34
1005 Lausanne | Switzerland

Visit us: www.frontiersin.org
Contact us: frontiersin.org/about/contact

REPRODUCIBILITY OF RESEARCH
Support open data and methods to enhance research reproducibility

DIGITAL PUBLISHING
Articles designed for optimal readership across devices

FOLLOW US
@frontiersin

IMPACT METRICS
Advanced article metrics track visibility across digital media

EXTENSIVE PROMOTION
Marketing and promotion of impactful research

LOOP RESEARCH NETWORK
Our network increases your article's readership