INTEGRATING PHILOSOPHICAL AND SCIENTIFIC APPROACHES IN CONSCIOUSNESS RESEARCH

EDITED BY: Christopher Gutland, Wenjing Cai and Anthony Vincent Fernandez
PUBLISHED IN: Frontiers in Psychology

frontiers Research Topics

frontiers

ISSN 1664-8714
ISBN 978-2-88966-925-7
DOI 10.3389/978-2-88966-925-7

About Frontiers

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

Frontiers Journal Series

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

Dedication to Quality

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

What are Frontiers Research Topics?

Frontiers Research Topics are very popular trademarks of the Frontiers Journals Series: they are collections of at least ten articles, all centered on a particular subject. With their unique mix of varied contributions from Original Research to Review Articles, Frontiers Research Topics unify the most influential researchers, the latest key findings and historical advances in a hot research area! Find out more on how to host your own Frontiers Research Topic or contribute to one as an author by contacting the Frontiers Editorial Office: frontiersin.org/about/contact

INTEGRATING PHILOSOPHICAL AND SCIENTIFIC APPROACHES IN CONSCIOUSNESS RESEARCH

Topic Editors:
Christopher Gutland, Sun Yat-sen University, China
Wenjing Cai, Shanghai Jiao Tong University, China
Anthony Vincent Fernandez, Kent State University, United States

Citation: Gutland, C., Cai, W., Fernandez, A. V., eds. (2021). Integrating Philosophical and Scientific Approaches in Consciousness Research. Lausanne: Frontiers Media SA. doi: 10.3389/978-2-88966-925-7

Table of Contents

frontiers in Psychology

EDITORIAL
published: 28 April 2021
doi: 10.3389/fpsyg.2021.683860

Editorial: Integrating Philosophical and Scientific Approaches in Consciousness Research

Christopher Gutland[1]*, *Wenjing Cai*[2] and *Anthony Vincent Fernandez*[3,4]

[1] Department of Philosophy, Center for Documentation and Research of Phenomenology, Sun Yat-sen University, Guangzhou, China, [2] School of Humanities, Shanghai Jiao Tong University, Shanghai, China, [3] Department of Philosophy, Kent State University, Kent, OH, United States, [4] Faculty of Philosophy, University of Oxford, Oxford, United Kingdom

Keywords: consciousness research, interdisciplinary, methodology, philosophy, science, phenomenology

Editorial on the Research Topic

Integrating Philosophical and Scientific Approaches in Consciousness Research

Consciousness has a long history as a topic of philosophical investigation. But its status as an object of scientific inquiry is a comparatively recent development. Consciousness seems to elude traditional approaches to scientific research because it is not directly accessible to third-person observational methods or objective systems of measurement. Researchers cope with this challenge in a variety of ways. It is common for scientists to appeal to philosophy to help them think through the nature of consciousness. Philosophical questions might inspire scientific studies and provide theoretical foundations for scientific research programs. And, with the rising interest in interdisciplinary research—including interdisciplinary collaboration—philosophers have likewise developed an interest in how, precisely, they might effectively integrate their conceptual or theoretical contributions with scientific approaches to the study of consciousness. With this Research Topic, we provide a platform for this integration of philosophical and scientific approaches in consciousness research.

Among the questions are: How can philosophical theories of consciousness generate testable hypotheses for scientific research? How should we go about formulating these hypotheses and designing studies that can test them? How can philosophy inform interview techniques that help us better understand the nature of conscious experience? And how can it guide mixed methods that bring together qualitative and quantitative techniques? How can philosophical accounts of consciousness that seem opposed to the assumptions of the natural sciences—such as transcendental theories of consciousness—be made compatible with current scientific research?

This Research Topic brings together a diverse range of articles that address these questions. We briefly introduce each article, proceeding from general methodological reflections to concrete applications and illustrations.

"A Map of Consciousness Studies: Questions and Approaches" by Niikawa is a second order investigation and multidimensional comparison of different theories encountered in consciousness research. It combines a top-down approach, listing and characterizing key questions in consciousness research, with a bottom-up approach regarding the different methodologies being applied. At the end, Niikawa uses the differentiations provided to compare the Integrated Information Theory and the Global Workspace Theory.

Edited and reviewed by:
Antonino Raffone, Sapienza University of Rome, Italy

***Correspondence:**
Christopher Gutland
christopher.gutland@gmx.net

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

Received: *22 March 2021*
Accepted: *30 March 2021*
Published: *28 April 2021*

Citation:
Gutland C, Cai W and Fernandez AV (2021) Editorial: Integrating Philosophical and Scientific Approaches in Consciousness Research. Front. Psychol. 12:683860. doi: 10.3389/fpsyg.2021.683860

Netland, in "The Living Transcendental—An Integrationist View of Naturalized Phenomenology," draws on Merleau-Ponty's phenomenology to outline a possible integration of the transcendental and the scientific. He characterizes transcendental consciousness as a structure of empirical nature. Through this integrationist view, a dialectical exchange between phenomenology and science can thus be made possible.

In "Ten Testable Properties of Consciousness," Tyler draws on the philosophy of Emergent Aspect Dualism to present properties of consciousness that are defined phenomenologically, yet testable on the level of neural substrates. Tyler bases this testability on a spatiotemporal isomorphism between the neural substrates and the experiential properties of consciousness, as well as the assumption that, in accordance with his understanding of emergence, the operative principles on the level of consciousness can be functionally dissociated from the levels below it. Tyler then presents 10 such properties and suggests how to test for them.

Keppler and Shani address the relation between the neural correlates of consciousness and its phenomenal, subjective aspect in their article "Cosmopsychism and Consciousness Research: A Fresh View on the Causal Mechanisms Underlying Phenomenal States." Cosmopsychism assumes the universe is imbued with a ubiquitous field of consciousness that has both an extrinsic physical appearance and an intrinsic manifestation which is phenomenological. The authors argue that when the neural cell assembly coherently oscillates, it acquires phenomenal properties by entering into a temporary liaison with the ubiquitous field of consciousness, whereby the zero-point field in quantum physics is the carrier of both primordial energy and consciousness.

In his article, "Phenomenological Skepticism Reconsidered: A Husserlian Answer to Dennett's Challenge," Belt reassesses Jean-Michel Roy's notion of phenomenological skepticism. Belt first describes Dennett's arguments against Husserlian phenomenology and presents versions of skepticism that they might imply. The major part of the article reconstructs key features of Husserl's methodology: époche and transcendental reduction, intentional analysis, eidetic variation and intersubjective validation. These reconstructions not only introduce and clarify these methodological facets. They also reveal how Dennett frequently misrepresents them and how they in fact serve to avoid many of the pitfalls Dennett mentions.

In their article, "The Hitchhiker's Guide to Neurophenomenology—The Case of Studying Self Boundaries with Meditators," Berkovich-Ohana et al. and her team provide a practical guide to Varela's Neurophenomenological Research Program. The authors outline various ways to bridge first- and third-person research using neurophenomenology and argue that researchers should alternate among the bridges depending on their experimental resources and domains of interest. Following this, they provide concrete examples of neurophenomenological studies from their own research program on the dissolution of self-boundaries in meditation.

In "A Phenomenological Paradigm for Empirical Research in Psychiatry and Psychology: Open Questions," Irarrázaval reflects on the suitability of qualitative research for the phenomenological study of psychopathology. She considers what makes an interview phenomenological, why we should conduct phenomenological interviews with patients, and how to perform an analysis of the data provided by such interviews, among other questions. In closing, she considers the issue of reality's mind-independence within the context of phenomenological research.

The article, "Framing a Phenomenological Mixed Method: From Inspiration to Guidance," contributed by Martiny et al., aims to integrate qualitative and quantitative methods in a framework of phenomenological mixed methods. By looking into several existing cases that apply mixed methods to the investigation of concrete social phenomena, the article provides guidance on how phenomenology can inform data generation, analysis, and interpretation.

In "The Problem of the Task. Pseudo-Interactivity as an Experimental Paradigm of Phenomenological Psychology," Wendt stresses the distinction between task and problem while simulating complex social interactions. Reassessing the psychology of thought, Wendt reveals how it offers a means to distinguish between the subject's experiential conceptualization of the task and her motivation to make it her problem. Wendt then introduces pseudo-interactivity, which draws on the psychology of thought and phenomenology, to investigate solely the experiential conditions of the situation in which a subject partakes in a problem-solving task. At the end, Wendt discusses an experiment simulating a pseudo-interactive semantically complex personal interaction.

To counter the denial of a true self, Sparby et al. combine psychology, phenomenology, and narratology in their article "The True Self. Critique, Nature, and Method." After illustrating the widespread belief in a true self in today's folk psychology and in historical traditions, the authors ask, "Could this belief not be grounded in reality?" Distinguishing between a thin and thick conception of the true self, they offer a defense against the claim that the true self is radically subjective and scientifically unverifiable. Lastly, the authors suggest a method for investigating the true self.

The article, "Inhibited Intentionality: On Possible Self-Understanding in Cases of Weak Agency," by Ingerslev, studies consciousness in unreflective actions and illuminates an inhibited form of intentionality where weak agency is involved. The article, drawing on phenomenological as well as psychoanalytical insights, proposes a diachronic account of consciousness that can make sense of the possible self-understanding of weak agency in terms of a process of appropriation of one's own action.

Strappini et al., in their contribution, "Empirical Evidence for Intraspecific Multiple Realization?" offer a definition of a psychological property of restored object identification and along this line examine case studies of visually impaired patients to support the Multiple Realization Thesis. This thesis is an anti-reductionist approach claiming both that mental properties are multiply realized and that mental processes can be implemented by different neural correlates.

When considered together, the articles within this Research Topic illustrate how philosophical and scientific approaches can be combined, both in principle and in concrete applications, and how this integration can advance consciousness research.

AUTHOR CONTRIBUTIONS

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

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: 22 October 2019
doi: 10.3389/fpsyg.2019.02250

The True Self. Critique, Nature, and Method

Terje Sparby[1,2], Friedrich Edelhäuser[2,3,4] and Ulrich W. Weger[1,2]*

[1]Department of Psychology and Psychotherapy, University of Witten/Herdecke, Witten, Germany, [2]Department of Psychology, Integrated Curriculum for Anthroposophic Psychology (ICURAP), University of Witten/Herdecke, Witten, Germany, [3]Department of Medicine, Integrated Curriculum for Anthroposophic Medicine (ICURAM), Witten/Herdecke University, Witten, Germany, [4]Department of Medicine, Institute of Integrative Medicine, Witten/Herdecke University, Witten, Germany

The history of philosophy gives us many different accounts of a true self, connecting it to the essence of what a person is, the notion of conscience, and the ideal human being. Some proponents of the true self can also be found within psychology, but its existence is mostly rejected. Many psychological studies, however, have shown that people commonly believe in the existence of a true self. Although folk psychology often includes a belief in a true self, its existence is disputed by psychological science. Here, we consider the critique raised by Strohminger et al., stating that the true self is (1) radically subjective and (2) not observable, hence cannot be studied scientifically (Strohminger et al., 2017). Upon closer investigation, the argument that the self is radically subjective is not convincing. Furthermore, rather than accepting that the true self cannot be studied scientifically, we ask: What would a science have to look like to be able to study the true self? In order to answer this question, we outline the conceptual nature of the true self, which involves phenomenological and narrative aspects in addition to psychological dimensions. These aspects together suggest a method through which this concept can be investigated from the first-person perspective. On a whole, we propose an integrative approach to understanding and investigating the true self.

Keywords: the true self, the self, first-person methods, consciousness, phenomenology

OPEN ACCESS

Edited by:
Christopher Gutland,
Sun Yat-sen University, China

Reviewed by:
Brentyn John Ramm,
Independent Researcher, Fremantle, Australia
James Morley,
Ramapo College, United States

****Correspondence:***
Terje Sparby
terje.sparby@uni-wh.de

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

Received: *13 May 2019*
Accepted: *19 September 2019*
Published: *22 October 2019*

Citation:
Sparby T, Edelhäuser F and Weger UW (2019) The True Self. Critique, Nature, and Method. Front. Psychol. 10:2250. doi: 10.3389/fpsyg.2019.02250

INTRODUCTION

Let us start with a quote: "Many people like to think they have an inner "true" self. Most social scientists are skeptical of such notions. If the inner self is different from the way the person acts all the time, why is the inner one the "true" self?" (Baumeister and Bushman, 2013, p. 75). This is how the notion of a true self is introduced in a recent textbook on social science. It is suggested that there is a conflict between folk psychology and science, where the true self is a notion that does not hold up to closer scrutiny. This view has recently been reinforced by a number of studies conducted by Strohminger and Nichols (2014) and Strohminger et al. (2017), showing that belief in a true self is indeed common while questioning its actual existence. Is the view we have of our "true self" merely a reflection of the socio-cultural environment in which we exist? And can someone have a "true self" that is good, even if they continually act in ways that are harmful?

Positing a chimera of an inherently good "true self," existing so deeply within the structure of someone's psyche that it may never make an appearance in reality may seem completely unwarranted. Not only does this put the true self beyond scientific observation, it also makes it seem like a hopelessly optimistic dream. Hence, although it is empirically clear that people make

use of the concept of a true self – in the sense of that which cannot change without someone becoming less of what they really are – there are weighty reasons to doubt whether the true self exists beyond the widespread belief in it. Since this belief is so common, could it be that it is in fact grounded in reality?

This is the question that we will explore in the following, outlining not only a suggestion for what the structure of the true self might be but also sketching out a method for investigating it. In doing this, we will also provide counter-arguments to the critique of the aforementioned true self. In our view, the true self can be viewed as having a kind of spiritual existence. It can appear in time but also exists beyond time. It may even be absent at different moments in time without ceasing to exist. Complete absence of the true self would, however, make it impossible to investigate. We take it that we are dealing with an essence of the Hegelian kind, i.e. an essence, the essence of which is to appear (and indeed, can there be an essence that never appears?). In other words, the true self cannot be so chimeric as to never enter the stage of actual life. However, such an object of study cannot be investigated adequately using conventional philosophical or psychological methods alone. We propose that the true self may be approached through a first-person method combining both philosophical reflection and introspective observation, as we will outline in section "Outline of a Comprehensive First-Person Method for Studying the True Self." Before introducing this method, we will look into the history and nature of the self and the true self in philosophy and psychology (section "Introduction"). This will follow with a response to critiques of the true self (section "The Problem of Radical Subjectivity and Observability of the True Self").

A Short Historical Account of the Self and the True Self

The self, one of the most central as well as critically discussed concepts in philosophy and psychology, has a long history. The idea that one has an underlying self in addition to a surface personality can be traced back to the notion that one has a soul that is potentially immortal. In the Egyptian culture, only the Pharaoh possessed an immortal, divine soul (akh) while alive. Only at the moment of death could other Egyptians gain such a soul (Waage, 2008). In Ancient Greek culture, Socrates was known for having heard an inner voice that indicated to him what he should (*Memorabilia* 1.1.4, 4.3.12, 4.8.1, *Apology* 12) and should not do (*Apology* 31c-d, 40a-b, *Euthydemus* 272e-273a). This was part of what lead to his demise, as he was accused of following other gods. The inner voice was a *daimonion*, a divine being (particular) to Socrates and not one of the gods condoned by the Athenian city-state. Such a private divine being is now commonly understood to refer to conscience in the Christian tradition (Schinkel, 2007, p. 97), which is connected to the moral essence – the true self – of an individual. The idea of a person's moral essence was developed further in Greek thought. For example, it was connected to the performance of specific virtues by Aristotle. Aristotle also suggested that "the true self of each" is the divine intellect or *nous* (NE, 1178, a2).

However, when answering the question "who are you?", it was for a long time customary to name one's ancestors. In ancient Rome, the firstborn son was the property of the *pater familias* until the death of the father. During the funeral procession, the son wore the father's death mask (Salemonsen, 2005). It may be noted that the word "mask" (*lat.* persona) is related to the word "person," suggesting that we can take on different identities but also that there is an underlying essence. Augustus is known for writing the first autobiography, inaugurating a genre defined by the idea that certain events and thoughts are more important than others when seeking to understand who someone is. Arguably, the Judeo-Christian religions also contributed to the view that all human beings have a divine core, regardless of background: "There is neither Jew nor Greek, there is neither bond nor free, there is neither male nor female: for ye are all one in Christ Jesus" (Galatians 3:28). In the Renaissance, Pico della Mirandola emphasized the notion of agency in his "Oration on human dignity", making God exclaim that it is a matter of will whether the human being shall become animal or divine, mortal or immortal:

> I have placed you at the very center of the world, so that from that vantage point you may with greater ease glance round about you on all that the world contains. We have made you a creature neither of heaven nor of earth, neither mortal nor immortal, in order that you may, as the free and proud shaper of your own being, fashion yourself in the form you may prefer. It will be in your power to descend to the lower, brutish forms of life; you will be able, through your own decision, to rise again to the superior orders whose life is divine (della Mirandola, 1996, p. 7).

For Kant, the self is that which provides transcendental unity to our thoughts and perceptions, in short, to all our experiences (Kant, 1904). Although the self cannot be known as it is in itself, in Kantian ethics, the individual is fully autonomous, free, when it acts according to rational principles (Kant, 1968). The individual manifests the kingdom of heaven on earth to the extent that ethical principles are adhered to as if they were natural laws. As a reaction to this, some philosophers, such as Sartre, point out that this view disregards the communal and social aspects of the self as well as its individuality and authenticity (Sartre, 2014). Rejecting Sartre's notion of authenticity, Foucault denied that there is any self that is given to us; claiming that we should rather view the self as a work of art:

> I think that from the theoretical point of view, Sartre avoids the idea of the self as something that is given to us, but through the moral notion of authenticity, he turns back to the idea that we have to be ourselves – to be truly our true self. I think that the only acceptable practical consequence of what Sartre has said is to link his theoretical insight to the practice of creativity – and not that of authenticity. From the idea that the self is not given to us, I think that there is only one practical consequence: we have to create ourselves as a work of art (Focault, 1997, p. 262).

Foucault points out that Sartre's notion of authenticity reintroduces a given measure for someone's true self. Foucault thinks we should be more radical in our rejection of any given content or measure of what constitutes the true self. Any such content or measure we must create ourselves. One may remark that even creative acts contain an element of or at least relate to something given, for example an inspiration or a framework of understanding. The idea of creating a self does not need to be thought of as a pure/arbitrary invention of something incomprehensible. Creative acts may be understood instead as the encounter between something given and subjective energy. In part, the subject identifies with the given, subjects itself to it, and in part, the subject recognizes the given as itself.

If we pause and summarize here, we can see that there is a whole host of ideas connected to the self in the Western canon (for a discussion of self, no-self and true self accounts in Asian traditions, see Siderits et al., 2010):

1. The self is a kind of essence, substance, or a soul that may or may not survive death
2. The self is the voice of conscience, the source of moral or authentic action
3. The self is divine, possibly created by God
4. The self is related to the past, to ancestry, and outward identity such as one's work
5. The self has a story connected to it that can be represented in a biography
6. The self provides unity to cognition and experience
7. The self is a free, autonomous agent
8. The self is essentially connected to other human beings and culture
9. The self has to be created

As we can see from this short and non-exhaustive list, the self is complex and may be conceived in conflicting ways. For example: Is the self-created by God or the individual? Is the self completely autonomous or is it thoroughly culturally determined? Is the self an essence or is it a story? None of these are necessarily contradictory, but much work is required to flesh out a comprehensive conception of the self. Do all these characteristics have something in common? This question is not easy to answer. If we cannot find a common characteristic in all the different definitions, we may have to concede that the self is simply a name for a host of unrelated ideas or aspects of human existence. On a closer look, each item on the list can potentially be said to be the true self. Even one's outward identity could arguably be seen to constitute a true self. Imagine a *puer aeternus*, a Peter Pan-like existence: someone who is reluctant about identifying with anything at all, preferring to stay adolescent indefinitely. For such a person, actually identifying with something could be said to be a realization of their true self (their true self would not necessarily be the specific outward identity but could be manifested by taking on a concrete and not fantastical identity). There is one way of conceiving what the nature of the true self is, which we will elaborate in the following, that does not imply that we have to make a choice about which specific self represents the true one. This is the conception of the true self as a whole that unifies the different selves. Moreover, the true self can not only be viewed as a whole but also as the manifestation of a specific moral self that has grown out of the past. The true self, on this conception, has both distanced itself from the past and integrated it, moving toward an ideal that is in one sense given, internally and from the past, but in another sense must also be created, or is only just coming into existence from the future.

The True Self in Philosophy and Psychology

Although the existence of the self is controversial in philosophy (Metzinger, 2003; Siderits et al., 2010; Ganeri, 2012), there are a number of influential philosophers who claim that there at least a minimal or core self exists. Such a view can be found both among traditional thinkers, such as Descartes, Leibniz, Kant, Hegel, Husserl, etc., and contemporary ones (MacIntyre, 1981; Taylor, 2012; Zahavi, 2017). Charles Taylor has specifically addressed the notion of a true self in the context of a discussion of negative and positive freedom (Taylor, 1985; Sparby, 2014). Negative freedom is the idea that one can realize one's true self insofar as there are no external restrictions on the self (and perhaps no internal restrictions such as fear). But where does the understanding of what actually counts as being the true self come from? If it comes, for instance, from a totalitarian state, then the "true self" may indeed be a false self since someone other than the self determines it. Hence it would follow that actualizing a true self is typically seen to include self-determination. It could of course be that content of a state prescribed true self accords *by coincidence* with the true self recognized by a person. This would not stop the person from actualizing the true self as long as the recognition is internally constituted through reflection and moral deliberation. However, if someone can determine themselves radically, does this not mean that the content of the true self is arbitrary? We believe that such problems can be solved with ideas as such "being-with-oneself in otherness" (Sparby, 2016). For example, acting according to one's true self does not exclude acting according to principles as long as these principles are recognized as stemming from the true self. Finding one's true self may involve finding oneself in another person, community, culture, etc. This does not mean that the true self is simply something given. Even creative processes can involve something approaching the self "from the outside", such as an inspiration. Again, the true self can be viewed as a whole, as something transcending the subject-object dichotomy, allowing for such events where something comes to the self seemingly from an external source (e.g. the voice of conscience), a source which is, however, more adequately conceived of as belonging to the self in a deeper, higher or more inclusive sense. It is of course possible that the voice of conscience might be an expression of an internalized dogmatic morality. However, this does not make it unreliable in principle. It means that what it dictates has to be viewed in light of an investigation of what its source might be, considering cultural factors specifically.

Is a person always acting in accordance with their true self if they act according to their self? The problem here is that the self is not only multifaceted but also contradictory given that different aspects are in conflict with each other. For example,

the human being can act out of principle or according to their desires. Both may be viewed or at least experienced as essential parts of one's identity, although these parts do not always harmonize. If one acts according to one's desire, another desire may not be fulfilled. If one acts morally, desires may fail to be satisfied at all. If one acts in a case where there is a moral dilemma, the true self seems to be constituted by that act. But what if I act based on wrong information, inherited cultural views, or delusion? Indeed, as we shall see, one of the main critiques of the true self is its radical subjectivity. The beliefs and actions that we ascribe to the true self depend on our worldview that is ultimately a reflection of the culture we belong to.

The field of psychology has contributed to our understanding of the self by gathering empirical support for the view that we are indeed ruled by external forces, such as unconscious desires, bias, and social conditioning. It has been shown that the experience of living a meaningful life is associated with having cognitive access to one's true self, and yet psychological research remains either skeptical or agnostic about the existence of it (Schlegel et al., 2013) despite the belief in a true self seems to be independent of personality type and culture (De Freitas et al., 2018). However, one can indeed find representatives of notions of a true self also in psychology. The true self is sometimes referred to as the I-self or self-as-process as opposed to the me-self or self-as-object (Ryan and Rigby, 2015). The former "concerns the conceptions, images, roles, statuses, and attributes associated with an identity," while the latter "concerns the inherent integrative tendencies of people to understand, grow, and create coherence in their experiences" (Ryan and Rigby, 2015, p. 246). The psychoanalyst Winnicott made explicit use of the concept of a true self, contrasting it with the false self (Winnicott, 1965). His view of the true self can be summarized as the self that is spontaneous, alive, and creative – the false self would then be a persona that lacks those characteristics (Rubin, 1998, p. 102). Numerous other terms are used for the true self such as the real self, the ideal self, the authentic self, the intrinsic self, the essential self, and the deep self [see overview of sources in Strohminger et al. (2017)]. Strohminger et al. have shown that people on average understand moral traits to be most fundamental to a person in addition to personality, memories, and desires, while characteristics related to perceptual abilities (e.g. near-sightedness) and psychical traits are perceived as having the least impact on who someone essentially is (Strohminger and Nichols, 2014). The essential differences between the self and the true self according to Strohminger et al. are that the self (1) encompasses the entire range of personal features, (2) is valence independent (it is inherently neither good nor bad) but (3) is perspective (first- or third-person) dependent, and (4) is cross-culturally variable, while the true self has an emphasis on (1) moral features, is (2) valence-dependent or positive by default, (3) perspective independent, and (4) cross-culturally stable (Strohminger et al., 2017, p. 3).

Strohminger et al. have also provided a particularly powerful formulation of the argument against the true self, which is quoted in full since it is the critique used as the background to our suggestion of what the nature of the true self is and how it can be studied:

> Is the true self also a scientific concept, one that can be used to describe how the mind actually works? Is there, in other words, a true self? The evidence reviewed here points to two properties relevant to this question. One: the true self depends on the values of the observer. If someone thinks homosexual urges are wrong, she will say the desire to resist such urges represents the true self (Newman et al., 2014). And if she scores high in psychopathy, she will assign less weight to moral features in her conceptualization of personal identity (Strohminger and Nichols, 2014). What counts as part of the true self is subjective, and strongly tied to what each individual person herself most prizes.
>
> Two: The true self is, shall we say, evidence-insensitive. Resplendent as the true self is, it is also a bashful thing. Yet people have little trouble imbuing it with a host of hidden properties. Indeed, claims made on its behalf may completely contradict all available data, as when the hopelessly miserable and knavish are nonetheless deemed good "deep down". The true self is posited rather than observed. It is a hopeful phantasm.
>
> These two features—radical subjectivity and unverifiability—prevent the true self from being a scientific concept. The notion that there are especially authentic parts of the self, and that these parts can remain cloaked from view indefinitely, borders on the superstitious. This is not to say that lay belief in a true self is dysfunctional. Perhaps it is a useful fiction, akin to certain phenomena in religious cognition and decision-making (Gigerenzer and Todd, 1999; Boyer, 2001). But, in our view, it is a fiction nonetheless (Strohminger et al., 2017, p. 7).

To reiterate, the problem facing the true self-view is that it is a conception tied to the values of a person, which are determined subjectively according to the structure of their personality, and by the culture and social environment in which that person exists. What the authors mean by "radical subjectivity" is, however, not clear. Does it mean that the values that a person uses to measure whether they live up to their true self are arbitrary, that the true self is based on a radical existential choice not grounded in anything, or that it is determined by biological, cultural, or social factors that happen to affect the person? These are issues that need to be untangled and answered. Furthermore, a good response is needed when arguing that the true self is not observable and therefore fictional. In particular, does it make sense to speak of a true self if that self never manifests? Can a person be called inherently good if they commit heinous crimes and consistently act in ways that are harmful to others, taking pleasure in their suffering?

In order to argue in favor of the existence of the true self, one must address the critique that it is a radically subjective notion and that it is unverifiable. Since we take the view that the self is not a thing with clearly defined borders but rather an organizing principle of a continual process, speaking of

"the existence" of a true self can be misleading. Nevertheless, one may claim that there is such an organizing principle and that the true self is neither radically subjective nor unverifiable. Before turning to that, we will provide a preliminary delineation of the true self that we will flesh out as we address the critique above.

A Thin and Thick Conception of the True Self and Their Unity

Two conceptions of the true self are implicit in what has been said above, which we will refer to as the thin and thick conception of the true self. One way to characterize them is to say that the thin conception is static: unchangeable, timeless, always the same. The thick conception is dynamic: developing, spread out over long changes of time, and continually emerging. The current objective in the following is to unite these two conceptions (in fact, to show how they are interdependent) and to investigate how such an account may be able to respond to the critique raised against the true self that we will focus on in section "The Problem of Radical Subjectivity and Observability of the True Self".

The thin conception of the true self is the idea that the self has a deeper and more essential nature; the true self is identical to this essential part of the self. Some of the properties attached to the self are accidental while others are essential. Someone can change their job and although they may have identified with their job, they do not really cease to be who they truly are when they change jobs. The true self as the essential self can consist of either one essential property or a set of properties. Sometimes, this is also referred to as the minimal self, which can be defined as the simple quality of subjective experience; the most fundamental experience of what it is like to be this or that subject (Zahavi, 2017). However, as pointed out by Fasching, the essential self's nature may be exactly a *bare existence*; not recognizable by any property. It simply is and we know it as something that can identify itself with potentially anything but can never be reduced to any specific property (Fasching, 2016). A similar view is presented by Ramm, who, using first-person experiments, argues that the self in itself both lacks sensory qualities and is single (Ramm, 2017).

If we conceive of the true self along these lines, the result would be rather indeterminate. There would be nothing more to it than what is common to all other selves: a simple and unique existence potentially aware of itself as such. Any identification of the self with a particular property, such as being a human, acting morally, or having been born in a certain place, would be fully irrelevant to the true self. But this seems wrong – or at least too indeterminate. Not only would it be at odds with typical conceptions of the true self, it would also conceptualize the true self in the form of a ghost with no bearing on its environment. This leads us to the thick conception of the true self [compare Galen Strawson's conception of the self, which differentiates between the self as a distinct mental entity and a subject of experience and the self as an agent, personality and diachronic continuity (Strawson, 1997)]. The thick conception of the true self connects it to certain substantial and moral properties such as being able to form memories or making an existential choice. Hence the thick conception where the true self consists of more determinate characteristics than bare existence is in accordance with how the true self is typically conceived in folk psychology. Is there a specific property or set of properties the self can identify with to become a true self or at least a "truer" self? Can one make a choice or live in a way that does not represent the ideal version of that individual? This certainly seems to be the case. But what is the measure according to which an act or a way of life can be judged as being in accordance with one's true self? Who or what decides what counts as a proper measure? What is it based on? Where does the true self come from? It will later be discussed how the true self is essentially related to both the past and the future. It will also be suggested that a certain conception of the true self can unite both the thin and thick version of it. Before turning to that, however, we turn to some discussions surrounding the true self in philosophy and psychology.

THE PROBLEM OF RADICAL SUBJECTIVITY AND OBSERVABILITY OF THE TRUE SELF

Here we will consider the two problems connected to the idea of the true self as identified by Nichols et al. above.

Radical Subjectivity

As we have seen, the problem of radical subjectivity relates to the notion that how someone conceives of their true self is dependent on what values they have. As we have stated earlier, there are more ways of interpreting what the claim that the true self is "radically subjective" means. It can mean that the true self is based on: (1) something completely arbitrary, (2) an ungrounded existential choice or (3) external factors, such as culture and biology. Although Strohminger et al. do not state explicitly which interpretation they have in mind, we think, based on the examples they give (sexual preference and psychopathy), that the third option is more likely. A person's sexual preference is rarely considered to be a choice but is rather understood to be based on biology and culture; psychopathy is hardly conceivable as a choice, but, again, is widely believed to be contingent upon biological, cultural or other environmental factors.

This, however, may seem surprising: Does not "radical subjectivity" mean something that involves arbitrariness or some form of creative or spontaneous choice? Since Strohminger et al. speak of the "radical subjectivity" of the true self as tied to what someone prizes or values, there might be some merit to the interpretation of it as being indeterminate in some way (not based on factors external to the self). But then again, the examples they mention point in another direction. So is the critique of the "true self" as radically subjective based on (1) the idea that it is radically arbitrary, random or contingent (what someone *happens* to value) or (2) the idea that the external factors that a person has happened to be exposed to due to the geographical location of their life and their inheritance has determined what they value?

It is highly unlikely that someone would hold the view that what someone values is completely arbitrary, based on something akin to the random result of throwing dice. For example, we value food because of biological needs, friendship because of social needs, and certain ideas because we find them enlightening. However, when we are presented with a moral choice or dilemma or when we are challenged with coming up with a plan for our next steps in life, our choice might seem subjective in the sense that it is creative or ultimately relies on a decision. But if it is creative, this does not mean that it is arbitrary – as we argued above in relation to Foucault. And if it is ultimately based on a decision, this does not mean that we do not have good reasons for acting the way we do, although we might have reasons to act in other ways as well. So the choice itself might be spontaneous, but that does not mean that it is arbitrary in the sense of not being grounded in reasons. And insofar as it is not clear to us what reasons are the best when considering a moral dilemma or committing to a life path, we could regard the choice as creative – but again, such creativity does not have to be arbitrary. What we are left with is the notion that someone's idea of their true self is radically subjective because it is based on what they happen to value, which in turn is based on the features of their personality. We will consider this in more depth.

Depending on one's sexual preference or whether one has a personality disorder such as psychopathy, one may conceive of the core of one's personality differently. This boils down to a claim there are a variety of different conceptions of the self and that therefore how someone defines their true self is subjective. Such a view, however, fails to consider the possibility that one may be right or wrong about their true self. If there were a true self, it would indeed be possible to make such mistakes. We cannot take it for granted that there is no true self based simply on the fact that people value things differently and conceive of their true nature accordingly. Even if I value money and claim that I am affluent, I would be mistaken about this claim if I have no money. Even though people value things differently, and the specific values someone has influence how they conceive of their essential nature, it does not follow that one's true self is merely an extension of what one happens to value.

However, it is still a significant point that one's conception of oneself tends to co-vary with one's cultural background. Could it not be the case that someone's true self harmonizes with what a specific culture dictates, while someone else in that culture could have a completely different true self; one that runs counter to the common views and values? How would someone know if they were mistaken, i.e. simply influenced by their culture, when it comes to viewing what their true identity is? The true self may indeed be fully individual. How does one uncover it? Perhaps, this is possible exactly by making mistakes or taking on or trying out identities that are not in accordance with one's true nature.

It seems strange or even wrong to say that by changing one's identity or taking on a different role, one suddenly lives according to one's true self. This indeed identifies the true self with the me-self – the true self would be a specific role, identity, job, etc. – which seems counterintuitive; should the true self not be a deep self, the self-as-process? If I change my identity and consider the new identity my true self, it implies that the former identity was a false self. But was it not the case that one aspect of the true self is exactly an underlying identity, one that cannot change simply by changing from one's surface identity to another? Without such an underlying identity it would not make sense to say that the former identity was a false self, because there is nothing to connect the two identities.

Indeed, the true self may be conceived of as that which unifies different conceptions of the more concrete selves (the me-selves) through a narrative (Polkinghorne, 1991; Gallagher, 2000; Schechtman, 2011), where the variations and mistakes are not necessarily plain errors, but rather essential parts of the process. By manifesting a unity within the different conceptions of the me-self, the true self is also manifested. This manifestation is not necessarily tied to a specific identify, a me-self, being right or wrong, true or false. The measure of the degree of manifestation is the degree of unity created by the processual self-conception. Since the self is also influenced and potentially challenged by different cultures, ethical norms, and worldviews, the unity increases to the extent the different cultures are encompassed, i.e. to the extent that difference is recognized and integrated in the true self.

This capacity of unity may manifest in different ways for different aspects of the true self. Take for example the ethical self, which as pointed out previously, is considered by many to be the true self. Even if one considers the true self to be the ethical self, it does not follow that the true self is radically subjective. What I value may be dependent on a whole range of factors, but that does not mean that the values cannot be judged objectively. There is a long tradition of discussion surrounding the question of whether ethics is objective. However, since there is no consensus on this issue, one cannot say with confidence that values are subjective. Does this mean that the true self is identical to a specific moral set of beliefs? Here, it is helpful to differentiate between different potential layers of the true (ethical) self: (1) the capacity of moral deliberation and action, (2) specific moral views, (3) individual moral or existential choices. At the most fundamental level, a moral self does not consist of a specific set of moral principles and beliefs, but rather of the capacity of ethics, i.e. the capacity of ethical deliberation. Even if one is mistaken about a specific ethical act, the capacity to deliberate offers continuity to the true self. Recognizing that a previous act is wrong is inherently a deepening of the capacity of morality. However, certain acts do not necessarily involve a universal ethical requirement; ethical individualism allows for certain acts being ethical measured only according to the individual (Hegge, 1988). Depending on talent and interest for example might be right for one person to pursue a life as an artist, while wrong for someone else. Furthermore, there may be both general and individual patterns of ethical development that needs to be taken into account. The unity of such patterns, the connection between good and bad actions, failure and success – like the inner coherence of a drama – would be what the true self is.

Evidence-Insensitivity

Let us look at the argument against observability again:

> The true self is, shall we say, evidence-insensitive. Resplendent as the true self is, it is also a bashful thing. Yet people have little trouble imbuing it with a host of hidden properties. Indeed, claims made on its behalf may completely contradict all available data, as when the hopelessly miserable and knavish are nonetheless deemed good 'deep down'. The true self is posited rather than observed. It is a hopeful phantasm. […] The notion that there are especially authentic parts of the self, and that these parts can remain cloaked from view indefinitely, borders on the superstitious (Strohminger et al., 2017).

There are two related but not identical claims that seem to be inherent in this argument: One is that the true self is in principle unobservable and hence it is an unscientific (superstitious) concept. The other is that what the true self cannot be revised based on evidence, removing it from the domain of science. Both claims will be addressed in the following.

The fact that some properties may be hidden does not in and of itself make the object connected to those properties in principle unavailable to science. Indeed, scientific activity consists of making what is hidden visible, for instance through inventions such as the microscope. However, basing the argument on a contrast hidden/visible implicitly limits the range of inquiry to what we can and cannot *see*, which is unwarranted. Some phenomena, specifically those that unfold in time, are indeed *constitutively dependent on some related properties being unavailable* ("hidden") *as the phenomenon manifests*. When a phenomenon manifests, something in the previous stage must be removed for a new stage to replace it. In other words, for something to manifest, something that once was, now has to be "hidden." For someone to say "the true self is not observable," for example, requires the word "the" to not appear (sound) when "true" is said. In fact, all other words must be "hidden" as well. What is consistent throughout the sentence is the invoked meaning. The meaning is partially invoked by each word and only fully invoked by the whole sentence (which cannot be present as a single instance in time, though perhaps as the retained meaning, something that includes the words and their sequence in a kind of concrete universal, i.e., a concept that is a whole containing its parts in it). Studying time-phenomena such as the self hence requires different methods than those that try to find and measure it at a specific moment in time. The latter approach may find it but only parts of it. Only a narrative that takes the whole into account can be an adequate method for studying the diachronic aspects of the self.

The claim that someone is "good deep down" despite all the evidence to the contrary is harder to counter. A "good" friend who never supports their friends is not a true friend. But is there any point at which someone loses the *capacity* for acting morally or being a good friend? Losing this capacity would also imply a loss of agency and the eligibility to be blamed. The self would be gone or at least not manifest in a basic sense. How could someone therefore provide evidence that the capacity really is absent? If someone always acted in a morally blameworthy way, what we could say, scientifically, i.e. based on observation, is that this person's true self is evil. However, a single good action would disprove that we have identified an essence. And a case of a person who consistently acts in a morally reprehensible way is hard to conceive. Is it someone who always acts so as to inflict the most pain possible? Is it someone not capable of any form of co-operation? Such a person would seem more like a machine than a human being. Even if we could conceive of such a person, we see no reason to reject the metaphysical possibility that such a person may change their ways. Maybe it would be possible to argue for the existence of evil true selves. Such an argument could very well be interesting but we suggest that for most persons it is possible to discover at least small acts of kindness, which would go to show the presence of a capacity for good. People who have indeed acted in reprehensible or in problematic ways and have changed provide a special area of study in relation to the true self. We take it for granted that such individuals exist. People who go through fundamental change toward good show that simple forms of observation and measurements at specific points in time are not adequate for studying the true self. An approach rather is required that takes long stretches of time into account. Given that there is a capacity for good or at least basic agency, a view that does not take this into account would be less truthful, i.e., less scientific, than a view that does take this into account.

It is still problematic that just as one can always correctly posit the capacity for good, one can also posit the capacity for evil. What is actually representative of one's true self then would seem to depend on what tendency manifests the most. For this reason, it seems appropriate to have a more abstract conception of the true self, i.e., as something that provides unity to life, and considers the relationship between good and bad acts. Actual human beings will probably never be so bad as to exclusively manifest evil actions and probably never be so good as to never do anything blameworthy. Considering what is good and bad, in the long run, requires historical perspectives. The scientific view is therefore also a view that is continually evolving with time.

One further objection to Strohminger et al. is that reliable methods already exist that objectively measure issues relating to the true self. For example, the ease with which people describe their true self is correlated with life meaning (Schlegel et al., 2009). However, such studies only assess the belief people have about the true self – not its existence. Strohminger et al.'s point is that the belief in the true self is evidence-insensitive in the sense that people are in principle unwilling to revise their view about what they believe their own or someone else's true self to be. The belief can only be confirmed, not rejected; hence, the true self is a non-scientific concept.

Another response to Strohimger et al.'s skepticism would follow a similar line of argument as Zahavi's response to Metzinger's claim that the self is an illusion or a model created by the brain. Zahavi's response is that the sense of self can be understood to constitute the self, or, in other words, the

existence of the self is nothing above and beyond the phenomenal experience of the self (Zahavi, 2005a). We would argue similarly in relation to the true self when faced with reductive arguments. The sense that people have of the existence of a true self can indeed be taken as constitutive of the true self. However, we wish to extend the concept of the true self to include specific life moments or developmental trajectories that manifest the true self, i.e., situations or ways of acting where the true self is not just a sense, but rather something that comes into existence. One could formulate this as an actualization of the potential true self. As we will outline in the next section, the sense of the true self extends not only into the past but also into the future. The existence of the true self in this way transcends time, although it can also appear or manifest for instance during significant life events – such as during Socrates' trial – where one's moral character is put to the test.

OUTLINE OF A COMPREHENSIVE FIRST-PERSON METHOD FOR STUDYING THE TRUE SELF

Understanding the true self as an activity in evolution and a process in metamorphosis involves conceptualizing it in a format that is most likely difficult to be nailed down with conventional, outwardly observable research methods. It is *per se* a first-person phenomenon and hence also requires a first-person mode of enquiry, although it also potentially involves behavioral aspects. It may manifest in a specific behavioral and even biological instantiation. This is however only the outer signature or correlate of the qualia of the phenomenon. This signature can be studied with conventional (e.g., behavioral or even physiological) research methods; the true self *in actu* as a first-person phenomenon, however, cannot be studied in this way. It can only indirectly be inferred from this signature mode of appearance. The approach outlined below can be seen as an extension of first-person approaches to the self that focus on its minimal, synchronous experience as presented by Ramm (2017), who involves for example directing attention to the point from which one looks at the world and investigating this point phenomenologically. The investigation reveals that this point has no visual features, but rather is transparent, single, etc. Through further experiments, one is lead to an experience of a minimal synchronic subject. The focus here is on diachronic aspects of the self, which are essential to study in order to develop a thick account of the self.

One can suspect that the true self can be grasped more by what it *can* become (causa finalis) rather than by what it *has* become (causa efficiens). This has further implications for the way it is studied. As an analogy, take the example of climate change. A small minority of people (mostly climatologists) made the earliest indications while researching subtle and even ambiguous symptoms of complex weather phenomena. For them to persist in their claim and stick to their account, they needed a good sense of trust in their reading and interpretation of the data and early indications. Moreover, they needed a vision of a future that might unfold if things continue in the manner they have developed so far. This was highly unusual and anachronistic at a time where climate change was still outside the conventional thinking style.

In a sense, the challenge that we see in this admittedly far-fetched example is somewhat related to the case of researching the true self. We need to investigate subtle and elusive symptoms to begin with and envision how this true self might unfold if given a chance to manifest and materialize within the constellation of potentials and situational factors with which the individual is endowed. Firmness is needed in envisioning the potentiality of this true self and a sense of trust that it can metamorphose from potential to reality. The moment it manifests as a reality in which to be studied in one way or another will have already crystallized into a given form to be considered a product rather than a process *in actu*. This would be an indication of a sub-component of the true self, not the true self proper.

The method we outline below is an extension of first-person approaches we have developed elsewhere, consisting of small groups of researchers investigating their experience through a series of meetings, note taking, comparison of results and repeated refinement of the experimental tasks that are carried out by the researchers themselves (Weger and Wagemann, 2015; Hackert et al., 2019). For more depth and precision of experiences involved in the descriptions of the events and tasks described below, micro-phenomenological interviews (Petitmengin, 2006) or self-inquiry can be employed.

We propose that a first-person method for studying the true self would include five steps:

1. The first step is developing a conceptual understanding of different possible accounts of what the true self might be. This involves envisioning possible worlds and future realities. Can a sense of one's true self be evoked through considering scenarios closer or further away from one's current life and identity? The sphere of the true self is not necessarily only that which is already instantiated but that which is still to come into existence. Without such conceptual guides, we are likely to miss the more subtle traces of the true self as mere background noise.
2. The second step is to consider significant life events (e.g., decisions, moral choices, challenging situations, illnesses, accidents, etc.) where one has the sense of either living up to or failing to live up to one's true self. Are there common markers of managing and failing to act in accordance with one's true self? What does the exploration of the sense of living in accordance with the true self reveal about the possible nature of one's own true self?
3. The third step is to consider the experience of the true self in the present moment. Which of my current properties and identities (gender, job, hair color, nationality, interests, philosophical outlook, etc.) relate to my true self? By employing a version of eidetic variation (Giorgi, 2009),

one can change any or all of these identities to see what can possibly be changed before the sense of who one is changes fundamentally. Furthermore, experiments such as described by Ramm (2017) can be employed to access the basic aspects of the synchronic subject. Is the sense of this subject similar to the sense of the true self one has developed to the present day? Additionally, different meditative techniques can be employed in order to heighten the awareness of the minimal self, for example by directing attention away from the awareness of specific thoughts, feelings, bodily sensations, the sense of self, and channeling them toward an awareness of awareness itself. To what extent is the true self connected to the minimal self and pure awareness and to what is it connected to specific properties of the actual/personal self? Can the true self be understood as integrating the minimal and personal self?

4. The fourth step is observing instances during the course of several weeks where one feels more vs. less at one with oneself. How do the instances where one feels more at one with oneself differ from those in which one feels less like oneself? How do such moments relate to the significant life events connected to the true self that were explored in step two?
5. The fifth step is "trusting" the true self into becoming – or one could also say: acting it out. This acting out has both a productive and a receptive side. The unfolding of the activity and getting to know it from inside inherently involve participating in its activity as well as cultivating a sense of receptivity for the inner echo that this activity produces. This fifth step is perhaps the most unusual form of scientific enquiry. This is a reminder that any form of research ultimately strives toward insight and the capacity for action (e.g., in the form of reproducing an effect that nature has created in a scientific physicochemical experiment).

Each step also involves checking one's understanding of the true self that was developed in the first step. Do any of the further steps lead to a deepening or change of one's initial conception of the true self? As such, this method involves both philosophical aspects as well as first-person experiments and first-person data gathered from memory. One could refer to such a method as "comprehensive" in that it involves investigating large developmental trajectories, present moment experiences, as well as how they relate to each other. It draws on different first-person methodologies that seem to be adequate for investigating the true self in the way we have presented in the previous sections. It may be noted that the method itself not necessarily presupposes any specific conception of the true self. It is therefore part of the method to reflect continually on what the true self means conceptually. Though the nature of the true self that we have suggested served as a guideline for developing the steps of the method outlined above, it may be that the actual first-person investigations of the true self following this method outlined here will lead to refinements both of the method as well as the account of the true self that we have argued for.

CONCLUSION

The basic function of the self is unity. It connects events in time and space into a single continuum of experience. To the extent that this unity is manifested, the true self is manifested. This can happen on different levels: (1) the core self – extending the continuity of the subjective sense of being – linking together orientation in space, time, and situation, and (2) the narrative self – creating unity throughout live events. Though we can say that there cannot be a narrative self without a core self, the converse is also true: The core self cannot actually exist – be aware of itself as a unity – without different moments in time being united within a time-structure. Hence, Zahavi is wrong, in part, in stating that:

> […]…it takes a self to experience one's life as a story. In order to begin a self-narrative, the narrator must be able to differentiate between self and non-self, must be able to self-attribute actions and experience agency, and must be able to refer to him- or herself by means of the first-person pronoun. All of this presupposes that the narrator is in possession of a first-person perspective (Zahavi, 2005b, p. 114).

Though this is half right, one can also say the opposite: There is no self without a minimal story, a beginning, middle, and end unfolding in time and united across time. However, it is also true that there needs to be an underlying self (unity) to the story. If no time has passed, it cannot be decided whether the self is indeed a self and hence the story/narrative and the minimal, phenomenological self are co-constitutive. In other words, the narrative and core self are co-constitutive and therefore inseparable. Although the latter may become ever more specified and deepened, this cannot happen without the core self. However, as the narrative self becomes more concrete in its various differentiations, the core self expands while not losing any of its being: It is that which is capable of being manifested as all the different concrete identities while not being fully identified with any single one of them. This self, a true self, can potentially be investigated following the methodical approach outlined above.

AUTHOR CONTRIBUTIONS

TS has written most of the manuscript. FE has taken part in conceptual development of the manuscript and commented on it. UW has taken part in the conceptual development of the manuscript, commented on it, and written parts of it.

FUNDING

This study received funding from the Software AG Stiftung. The funder was not involved in the study design, collection, analysis, interpretation of data, the writing of this article, or the decision to submit it for publication.

REFERENCES

Baumeister, R. F., and Bushman, B. J. (2013). *Social psychology and human nature*. Cengage Learning: Wadsworth.

Boyer, P. (2001). *Religion explained*. New York, NY: Basic Books.

De Freitas, J., Sarkissian, H., Newman, G. E., Grossmann, I., De Brigard, F., Luco, A., et al. (2018). Consistent belief in a good true self in misanthropes and three interdependent cultures. *Cogn. Sci.* 42(Suppl. 1), 134–160. doi: 10.1111/cogs.12505

della Mirandola, G. P. (1996). *Oration on the dignity of man*. Washington, DC: Regenery Publishing.

Fasching, W. (2016). The non-plurality of the I. on the question of the ultimate subject of experience. *J. Conscious. Stud.* 23, 140–157.

Focault, M. (1997). *Ethics: Subjectivity and truth*. New York: The New Press.

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

Ganeri, J. (2012). *The self: Naturalism, consciousness, and the first-person stance*. Oxford: Oxford University Press.

Gigerenzer, G., and Todd, P. M. (1999). *Simple heuristics that make us smart*. Oxford, England: Oxford University Press.

Giorgi, A. (2009). *The descriptive phenomenological method in psychology. A modified Husserlian approach*. Pittsburg: Duquensne University Press.

Hackert, B., Lumma, A. L., Menzel, P., Sparby, T., and Weger, U. (2019). Enquiring into the qualitative nature of anger: challenges and strengths of the introspective method. *Curr. Psychol.* doi: 10.1007/s12144-019-00221-0

Hegge, H. (1988). *Frihet, individualitet og samfunn. En moralfilosofisk, erkjennelsesteoretisk og sosialfilosofisk studie i menneskelig eksistens*. Oslo: Institutt for filosofi.

Kant, I. (1904). *Kritik der reinen Vernunft*. Berlin: Georg Reimer.

Kant, I. (1968). *Kritik der Praktischen Vernunft*. Berlin: De Gruyter.

MacIntyre, A. (1981). *After virtue*. Notre Dame, IN: University of Notre Dame Press.

Metzinger, T. (2003). *Being no one*. Cambridge, Mass: MIT Press.

Newman, G. E., Bloom, P., and Knobe, J. (2014). Value Judgments and the True Self. *Personal. Soc. Psychol. Bull.* doi: 10.1177/0146167213508791

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

Polkinghorne, D. E. (1991). Narrative and self-concept. *J. Narrat. Life Hist.* 1, 135–153. doi: 10.1075/jnlh.1.2-3.04nar

Ramm, B. (2017). Self-experience. *J. Conscious. Stud.* 24, 142–166.

Rubin, J. B. (1998). *A psychoanalysis for our time: Exploring the blindness of the seeing*. New York: New York University Press.

Ryan, R. M., and Rigby, C. S. (2015). "Did the Buddha have a self?: no-self, self, and mindfulness in Buddhist thought and Western psychologies" in *Handbook of mindfulness: Theory, research, and practice*. eds. K. W. Brown, J. D. Creswell, and R. M. Ryan (New York: The Guilford Press), 245–265. Available at: https://paloaltou.idm.oclc.org/login?url=http://search.ebscohost.com/login.aspx?direct=true&db=psyh&AN=2015-10563-014

Salemonsen, H. (2005). *Under Kunnskapens Tre*. Oslo: Vidarforlaget.

Sartre, J.-P. (2014). *Notebook for an ethics*. Chicago: University of Chicago Press.

Schechtman, M. (2011). "The narrative self" in *The Oxford handbook of the self*. ed. S. Gallagher (Oxford: Oxford University Press).

Schinkel, A. (2007). *Conscience and conscientious objections*. Amsterdan: Pallas Publications.

Schlegel, R. J., Hicks, J. A., Arndt, J., and King, L. A. (2009). Thine own self: true self-concept accessibility and meaning in life. *J. Pers. Soc. Psychol.* 2, 473–490. doi: 10.1037/a0014060

Schlegel, R. J., Smith, C. M., and Hirsch, K. A. (2013). "Examining the true self as a wellspring of meaning" in *The experience of meaning in life: Classical perspectives, emerging themes, and controversies*. eds. J. A. Hicks, and C. Routledge (Dordrecht: Springer Netherlands).

Siderits, M., Thompson, E., and Zahavi, D. (2010). *Self, no self?: Perspectives from analytical, phenomenological, and Indian traditions*. Oxford: Oxford University Press.

Sparby, T. (2014). "Hegel and the foundation of right" in *Contemporary philosophy: A new survey*. Vol. 12. ed. G. Fløistad (Dordrecht: Springer).

Sparby, T. (2016). Rudolf Steiners idea of freedom. As seen in the panorama of Hegel's dialectic. *Epoché* 21, 173–196. doi: 10.5840/epoche201682970

Strawson, G. (1997). The self. *J. Conscious. Stud.* 4, 405–428.

Strohminger, N., Knobe, J., and Newman, G. (2017). The true self: a psychological concept distinct from the self. *Perspect. Psychol. Sci.* 12, 551–560. doi: 10.1177/1745691616689495

Strohminger, N., and Nichols, S. (2014). The essential moral self. *Cognition*. doi: 10.1016/j.cognition.2013.12.005

Taylor, C. (1985). "What's wrong with negative liberty?" in *Philosophy and the human sciences. Philosophical papers 2*. ed. C. Taylor (Cambridge, MA: Cambridge University Press).

Taylor, C. (2012). *Sources of the self. The making of modern identity*. Cambridge, MA: Harvard University Press.

Waage, P. N. (2008). *Jeg: Individets Kulturhistorie*. Oslo: Schibsted.

Weger, U., and Wagemann, J. (2015). The challenges and opportunities of first-person inquiry in experimental psychology. *New Ideas Psychol.* 36, 38–49. doi: 10.1016/j.newideapsych.2014.09.001

Winnicott, D. W. (1965). *The maturational process and the facilitating environment. Studies in the theory of emotional development*. New York: International Universities Press, Inc.

Zahavi, D. (2005a). Being someone. *Psyche*. 11, 1–20.

Zahavi, D. (2005b). *Subjectivity and selfhood. Investigating the first-person perspective*. Cambridge, MA: MIT Press.

Zahavi, D. (2017). "Thin, thinner, thinnest: defining the minmal self" in *Embodiment, enaction, and culture: Investigating the constitution of the shared world*. eds. C. Durt, T. Fuchs, and C. Tewes, (Cambridge, Massachusetts: The MIT 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

PERSPECTIVE
published: 05 March 2020
doi: 10.3389/fpsyg.2020.00371

Cosmopsychism and Consciousness Research: A Fresh View on the Causal Mechanisms Underlying Phenomenal States

Joachim Keppler[1] *and Itay Shani*[2]*

[1] Department of Consciousness Research, DIWISS, Roth, Germany, [2] Department of Philosophy, Sun Yat-sen University, Zhuhai Campus, Zhuhai, China

Despite the progress made in studying the observable exteriors of conscious processes, which are reflected in the neural correlates of consciousness (NCC), there are still no satisfactory answers to two closely related core questions. These are the question of the origin of the subjective, phenomenal aspects of consciousness, and the question of the causal mechanisms underlying the generation of specific phenomenal states. In this article, we address these questions using a novel variant of cosmopsychism, a holistic form of panpsychism relying on the central idea that the universe is imbued with a ubiquitous field of consciousness (UFC). This field is understood as a foundational dual-aspect component of the cosmos, the extrinsic appearance of which is physical in nature and the intrinsic manifestation of which is phenomenological in nature. We argue that this approach brings a new perspective into play, according to which the organizational characteristics of the NCC are indicative of the brain's interaction with and modulation of the UFC. Key insights from modern physics suggest that the modulation mechanism is identical with the fundamental mechanism underlying quantum systems, resulting in the conclusion that a coherently oscillating neural cell assembly acquires phenomenal properties by tapping into the universal pool of phenomenal nuances predetermined by the UFC, or more specifically, by entering into a temporary liaison with the UFC and extracting a subset of phenomenal tones from the phenomenal color palette inherent in the basic structure of the UFC. This hypothesis is supported by a substantial body of empirical evidence.

Keywords: neural correlates of consciousness, hard problem of consciousness, explanatory gap, ubiquitous field of consciousness, zero-point field, modulation mechanism, quantum systems, long-range coherence

OPEN ACCESS

Edited by:
Wenjing Cai,
Shanghai Jiao Tong University, China

Reviewed by:
William Seager,
University of Toronto Scarborough, Canada
Gregory Miller,
University of Liverpool, United Kingdom

***Correspondence:**
Itay Shani
ishani479@hotmail.com

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

Received: *05 January 2020*
Accepted: *18 February 2020*
Published: *05 March 2020*

Citation:
Keppler J and Shani I (2020) Cosmopsychism and Consciousness Research: A Fresh View on the Causal Mechanisms Underlying Phenomenal States. Front. Psychol. 11:371. doi: 10.3389/fpsyg.2020.00371

CONSCIOUSNESS AND THE NCC

Neuroscientific approaches to the study of consciousness assign pride of place to the task of progressively charting and analyzing the *neural correlates of consciousness* (NCC), i.e., the neural mechanisms jointly sufficient for eliciting specific types of conscious experiences (Crick and Koch, 2003; Tononi and Koch, 2008). The search for the NCC is motivated, in large part, by the belief that they are *more* than mere correlates, namely, that such neural mechanisms provide the causal-material basis for consciousness. Significantly, many neuroscientists also hold (or have

held), optimistically, that an increased knowledge of the NCC will eventually shed light on the fundamental riddle known as the *hard problem* of consciousness, namely, the problem of understanding what it is about the brain which enables it to generate something as remarkable and unique as subjective phenomenal experience. In short, from the point of view of conventional neuroscientific lore the activity patterns constituting the NCC are not just observable concomitants of subjective experience in highly complex animals, but, rather, the ultimate foundation of consciousness (see, e.g., Seth et al., 2006).

In identifying the NCC with the ultimate basis of consciousness, this conventional approach is beset by two cardinal deficiencies. First, it severely restricts the spectrum of the possible causal mechanisms underlying consciousness, a restriction questionable on both empirical and theoretical grounds. Second, it remains orthogonal to the concerns driving the hard problem, unable to address these concerns head-on.

Before we attend to these problems, it is worth recalling, first, that few would deny that consciousness, as manifested in humans and in other advanced animals such as primates, dolphins, or birds (for example), bears special connection to the brain, and in particular to the specific processes and activity patterns which neuroscientists identify as the NCC. However, notwithstanding that there is a special connection between consciousness and the brain, and notwithstanding the relevant neuroscientific evidence, the precise *nature* of this connection remains an open question. In particular, the idea that consciousness is identical to such brain processes, or that these processes generate consciousness from utter insentience, is an *interpretation* of the data—it is something which neither the phenomenal, nor the behavioral, nor the neurophysiological data necessitate. To be sure, knowledge of the NCC should inform and constrain our efforts to understand consciousness and to shed light on the nature of the psychophysical nexus, but it does not deliver ready-made answers to the above-mentioned core questions.

AGAINST AN UNNECESSARY LIMITATION OF THEORETICAL HORIZONS

Consider now the restriction of the spectrum of possible causal mechanisms underlying macro-scale phenomenal consciousness. One sense in which cerebral chauvinism is ill-advised is evinced in the accumulation of evidence suggesting that the bounds of consciousness in the living world may far exceed cranial circumscription. To begin with, some highly intelligent creatures such as octopuses and other cephalopods are endowed with large neural ganglia on their arms, supporting sophisticated forms of sensing and moving with significant degree of autonomy from the octopus' brain (Hanlon and Messenger, 1996; Godfrey-Smith, 2013). More radically still, there is growing evidence for the existence of complex behavior in organisms lacking brains altogether. An intriguingly broad array of cognitive abilities is being progressively unveiled in simple eukaryotes, prokaryotes, and plants. Variegated forms of perception and behavioral plasticity, information processing, anticipation, memory, learning, valence, problem solving, communication, and cooperation are attributed to various brainless organisms from slime molds (Nakagaki et al., 2000; Reid et al., 2012), to bacteria (Ben-Jacob et al., 2006; Lyon, 2015), to plants (Trewavas, 2014; Gagliano, 2017).

In congruence with such studies, there is also a growing tendency to view neuronal networks as but one special case (albeit particularly powerful) of a general network dynamics whose fundamental principles are exemplified throughout the entire spectrum of biological life (Lyon, 2015; Baluška and Levin, 2016). In other words, many cognitive functions which in creatures such as Macaque monkeys, bees, or humans, are mediated through cerebral activity appear to be manifest, to some degree, in different forms of life (such as plants, slime molds, or bacteria) using alternative types of informational networks: be it methylation DNA networks, root systems, cytoskeletal elements, non-neural bioelectricity, calcium signaling, and so on.

While such studies often do not involve direct reference to consciousness (but see Trewavas, 2014; Baluška and Reber, 2019), the steady growth in evidence attesting to the existence of sophisticated cognitive repertoires throughout life's spectrum puts increasing pressure on the orthodox notion that consciousness and the NCC are coextensive. As soon as we cease taking such coextension for granted, we enjoy greater freedom to consider a wider range of possible causal mechanisms as potential candidates for a comprehensive explanation of consciousness.

MINDING THE EXPLANATORY GAP

Another cause for skepticism regarding the view that the NCC provide the ultimate basis of phenomenal experience is the familiar hard problem of consciousness (Chalmers, 1995). For given any set of neural configurations proposed as a proper physical underpinning for consciousness, there remains the question *why* such configurations should culminate in subjective experience. In the words of some notable early observers, the chasm between the physical and the phenomenal (as these are canonically understood by science, philosophy, and commonsense) appears to be "intellectually impassable" (Tyndall, 1879, 18), with the result being that the hypothesis that experience comes about through the irritation of nervous tissue "is just as unaccountable as the appearance of Djin when Aladdin rubbed his lamp" (Huxley, 1866, 193).

Underlying this gap is a fundamental dichotomy between the objective and the subjective. Science approaches its objects of study from an objective, third-person, perspective. Its descriptions are confined to the outward appearance of things, even when such "things" (objects, processes, events, or mechanisms) unfold inside the body or brain. It concerns itself exclusively with the behavior and structure of causal agents, that is, with the observable exteriors of its target explananda. In contrast, consciousness is a subjective, first-person, phenomenon. Its inner presence constitutes a manifest immanent reality irreducible to observable behavior and structure. Thus, in any approach confined to externals consciousness is bound to remain alien: identified, perhaps, but neither fully assimilated nor

properly explained. For this reason, it is imprudent to expect that more elaborate accounts of the neural basis of consciousness could ever be sufficient to address the challenge posed by the hard problem.

WEAVING SCIENCE AND PHILOSOPHY SYNERGETICALLY

What is needed, we submit, is a fresh outlook. Inclusive of consciousness and the intrinsic dimension of things but, at the same time, hospitable to objective findings and to rigorous scientifically based analysis. In this respect, the contribution of philosophy is vital. Philosophy's quest is maximally comprehensive in that it seeks to understand reality as a whole. As such, it must take into consideration the outer as well as the inner dimension of things. Moreover, its ultimate goal of arriving at an integral picture of reality in its entirety implies a commitment to strive to make sense of the connection between these two complementary aspects—the objective and the subjective. Consequently, it has ample historical and conceptual resources to draw upon in the effort to contribute to an improved understanding of the psychophysical nexus. In particular, we believe that some philosophical ideas recently rediscovered and redeveloped within the fields of metaphysics and the philosophy of mind give fresh impetus to consciousness research in that they provide a conceptual matrix opening up new interpretations of the neuroscientific body of evidence and, potentially, leading to unprecedented research strategies. In this spirit, we present the central ideas behind a novel variant of *cosmopsychism*, a holistic form of panpsychism from the genus of priority cosmopsychism that relies on the assumption of a cosmic level of consciousness serving as the ultimate bedrock of experiential reality (Keppler, 2012; Shani, 2015; Shani and Keppler, 2018; for different variants of priority cosmopsychism, see Mathews, 2011; Goff, 2017; Nagasawa and Wager, 2017).

THE CORNERSTONES OF A NOVEL COSMOPSYCHIST PARADIGM

Our approach is based on the central idea that the universe is imbued with an *inherently sentient medium*, hereafter referred to as *ubiquitous field of consciousness* (UFC). In order to avoid substance dualism, which would immediately present us with the problem of causal interaction, we require this field to be seamlessly embedded in the edifice of modern physics. Consequently, we posit that the UFC is a foundational *dual-aspect* component of the cosmos, the extrinsic appearance of which is physical in nature and the intrinsic manifestation of which is phenomenological in nature (more on this in the concluding section). As with all fields that play a role in physics, the extrinsic nature of the UFC should reveal itself in energetic form, which is reflected in a spectrum of oscillations, the so-called normal modes. Moreover, it is to be expected that in its initial state the field satisfies all essential symmetry requirements (isotropy, homogeneity, scale invariance, Lorentz invariance), entailing that there is no preferential direction, location, and reference system. This leads directly to the conception of the undisturbed UFC as an omnipresent, formless ocean of activity with completely uncorrelated modes (Keppler, 2016, 2018a). We then postulate that each normal mode is associated with an elementary phenomenal hue, implying that the entire phenomenal "color palette" is represented by the oscillatory spectrum of the UFC, with the terms borrowed from color vision being understood as illustrative metaphors that we use here and in the following in a broader sense to cover the entire range of phenomenal qualities. Accordingly, from the phenomenological point of view, the ground state of the UFC can be described as a shapeless, undifferentiated ocean of consciousness in the basic structure of which all shades of phenomenal awareness are inherent (Keppler, 2012; Shani and Keppler, 2018). Even though, from the perspective of our paradigm, no concrete conscious state can be assigned to the maximally disordered ground state of the UFC, experiences collected in deep states of meditation suggest that this ground state may be characterized as a maximally unified phenomenal state (for a more detailed discussion, see Shani and Keppler, 2018). The ubiquitous background field thus constitutes an entity that embodies the principles of physics and at the same time contains within itself the phenomenological basis of ultimate reality (see **Figure 1**).

Following this line of thought, it is natural to assume that conscious systems must be equipped with a fundamental mechanism by means of which they are able to influence the basic structure of the UFC, resulting in modified UFC states each of which displays complementary, intimately related physical and phenomenal properties. Without such a mechanism, there would be no conscious states other than the ground state of the UFC. Therefore, *a distinctive feature of conscious systems in comparison to non-conscious systems must be the capacity to modulate the omnipresent field of consciousness*, imposing constraints on the modulation mechanism. A look at the previously postulated properties of the UFC, according to which each normal mode of the field is linked to an elementary phenomenal hue, gives rise to the hypothesis that specific complex states of consciousness are formed by binding together specific sets of field modes. Consequently, using a rendering that is in accordance with modern physics, we argue that conscious states are caused by the dynamic interaction of a physical system with the UFC, provided that different modes, corresponding to the resonance frequencies of the system, are involved in this interaction and that the physical system establishes a transiently stable relationship with the background field resulting in the phase-locked coupling of the participating modes (Shani and Keppler, 2018). From this point of view, a physical system acquires phenomenal properties by entering into a temporary liaison with the cosmic field of consciousness and extracting a subset of phenomenal tones from the spectrum of all phenomenal tones potentially present in the field (see **Figure 2**). Conversely, this means that the phenomenal shades a given system can extract are determined by the range of dynamically stable states arising from the system-specific set of resonance frequencies. As a direct consequence, systems that cannot achieve dynamic equilibrium with the UFC have no access to its immanent phenomenal color palette and, hence, cannot generate phenomenal states (Shani and Keppler, 2018).

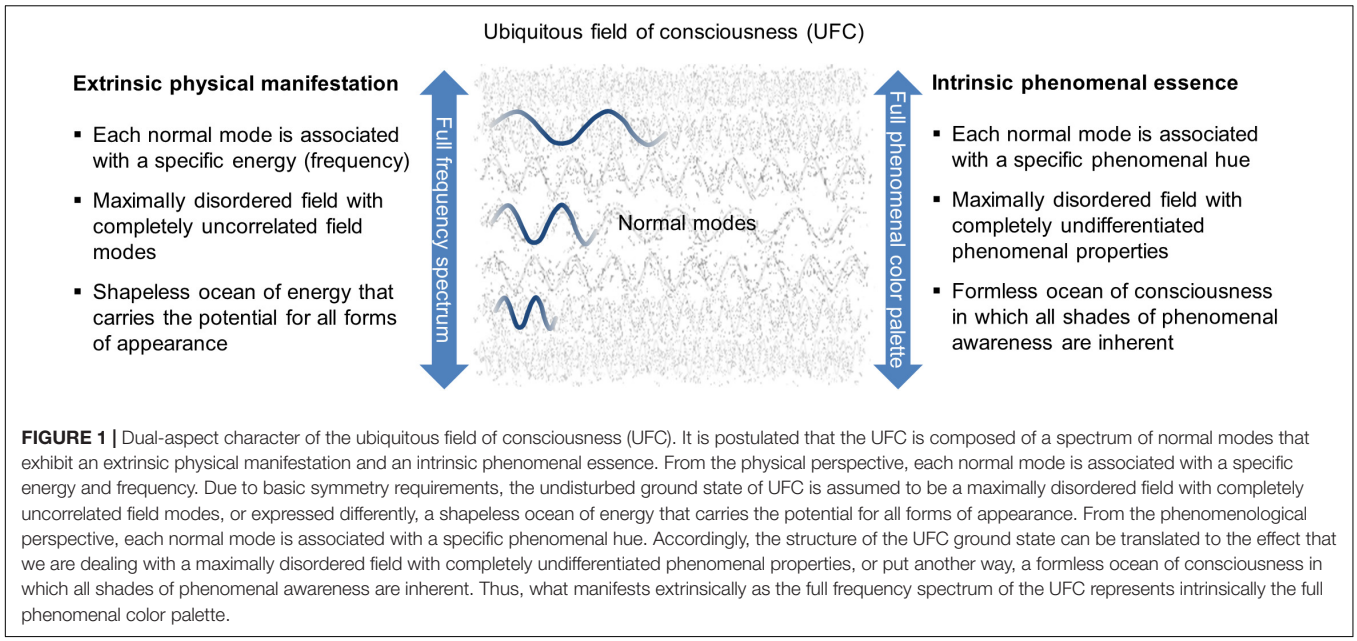


FIGURE 1 | Dual-aspect character of the ubiquitous field of consciousness (UFC). It is postulated that the UFC is composed of a spectrum of normal modes that exhibit an extrinsic physical manifestation and an intrinsic phenomenal essence. From the physical perspective, each normal mode is associated with a specific energy and frequency. Due to basic symmetry requirements, the undisturbed ground state of UFC is assumed to be a maximally disordered field with completely uncorrelated field modes, or expressed differently, a shapeless ocean of energy that carries the potential for all forms of appearance. From the phenomenological perspective, each normal mode is associated with a specific phenomenal hue. Accordingly, the structure of the UFC ground state can be translated to the effect that we are dealing with a maximally disordered field with completely undifferentiated phenomenal properties, or put another way, a formless ocean of consciousness in which all shades of phenomenal awareness are inherent. Thus, what manifests extrinsically as the full frequency spectrum of the UFC represents intrinsically the full phenomenal color palette.

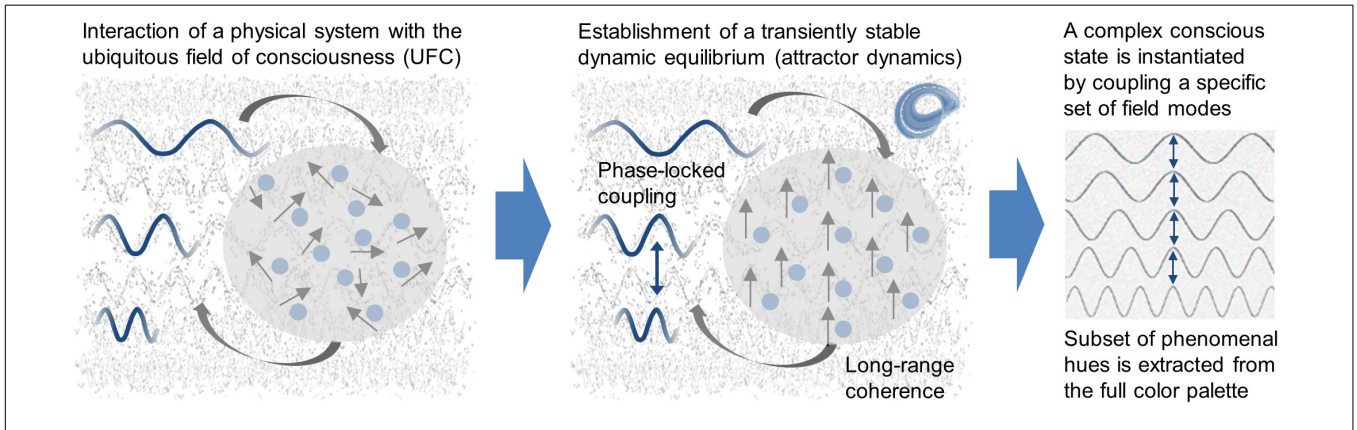


FIGURE 2 | Causal mechanism underlying phenomenal states. It is postulated that conscious systems must be equipped with a fundamental mechanism by means of which they are able to influence the basic structure of the ubiquitous field of consciousness (UFC). This requires the interaction of a physical system with the UFC in such a way that a transiently stable dynamic equilibrium, a so-called attractor state characterized by long-range coherence, is established in which the involved field modes enter into a phase-locked coupling. Based on the conception that each field mode is linked to an elementary phenomenal hue, the coupling of a specific set of field modes results in the extraction of a subset of phenomenal hues from the full phenomenal color palette and, hence, in the instantiation of a complex conscious state.

THE NOVEL COSMOPSYCHIST PARADIGM MEETS REALITY

It is now of vital significance that recent developments in modern physics are fully compatible with the above considerations. This applies particularly to stochastic electrodynamics (SED), "a branch of physics that affords a look behind the scenes of quantum mechanics and quantum field theory (QFT)," thereby "unveiling the mechanisms that account for the quantum behavior of matter" (Keppler, 2018a, 2). The foundations of SED, which date back to the 1960s, have been permanently refined in the endeavor to build a solid conceptual framework for quantum theory (Marshall, 1963, 1965; Boyer, 1969, 1975; De la Peña and Cetto, 1994, 1995, 1996, 2001, 2006; De la Peña et al., 2009, 2015).

A pivotal ingredient of this framework is "an all-pervasive electromagnetic background field, called zero-point field (ZPF), which, in its original form, is a homogeneous, isotropic, scale-invariant, and maximally disordered ocean of energy with completely uncorrelated field modes and a unique power spectral density" (Keppler, 2018a, 2). Based on this idea, the key findings of SED may be summarized to the effect that

1. "Every material system can be regarded as an open stochastic system in permanent contact with the random ZPF" (Keppler, 2018a, 2),
2. The dynamic interaction between a physical system and the ZPF can achieve energetic equilibrium, given that "the interaction strength between the oscillating components

and the relevant field modes, for which the system exhibits a strong resonant behavior, exceeds disturbing forces, such as thermal noise" (Shani and Keppler, 2018, 397),

3. "A system in equilibrium with the ZPF falls into a dynamically stable state, that is, an attractor, and displays *quantum behavior*" (Shani and Keppler, 2018, 397),
4. "The orchestration of an attractor requires the initially chaotic ZPF to change over to a partially ordered state that is characterized by an attractor-specific set of phase-locked field modes" (Keppler, 2018a, 2), which has the consequence that "all the components of the system are effectively coupled through the ZPF, giving rise to collective cooperation and *long-range coherence*" (Shani and Keppler, 2018, 398).

In light of these insights, "SED commends itself as a promising approach for the integration of consciousness into a coherent theoretical framework" (Keppler, 2016, 352). In particular, the findings listed above "suggest that the ZPF is perfectly suited for playing the dual role as the carrier of both primordial energy and consciousness" (Shani and Keppler, 2018, 399), which amounts to identifying the UFC with the ZPF. Moreover, in view of the previously formulated postulates relating to phenomenal states, the discoveries of SED support the assertion that "the mechanism underlying quantum systems has all the makings of a truly fundamental mechanism behind conscious systems, leading to the assumption that *conscious systems extract their phenomenal qualities from the phenomenal color palette immanent in the ZPF*" (Keppler, 2018b, 3). As a derivation, "conscious systems can be expected to display quantum behavior," meaning that "*the formation of transiently stable coherent states is an essential prerequisite for conscious awareness*" (Keppler, 2018b, 3).

Exactly this expectation is confirmed, especially as "the currently available body of evidence and the entirety of observations suggest that the brain has all the characteristics of a macroscopic quantum system" (Keppler, 2013, 3), which is substantiated in the following. Limiting ourselves for the moment to conscious perception, it is widely accepted "that the NCC are related to large-scale synchronization in the brain" (Keppler, 2018b, 3), a conclusion that is based on a considerable amount of neurophysiological data (Crick and Koch, 1990; Desmedt and Tomberg, 1994; Rodriguez et al., 1999; Engel and Singer, 2001; Melloni et al., 2007; Doesburg et al., 2009; Gaillard et al., 2009; Singer, 2015). A closer examination of the data (Freeman, 1991, 2004, 2005, 2007, 2009) reveals "that the NCC can be equated with attractors distinguishing themselves by a high degree of coherence between spatially distributed cortical areas and that our streams of conscious perception are based on the recurring formation and dissolution of such coherent states" (Keppler, 2018b, 3). These insights corroborate the assertion "that the NCC bear on quantum coherence since a rigorous description of the observed features, such as macroscopic pattern formation and second-order phase transitions, requires the formalism of quantum physics" (Keppler, 2018b, 3), which was also clearly emphasized by Freeman and Vitiello (2006, 2007). Including the previously enumerated findings of SED, this indicates "that the ZPF is involved in the orchestration of coherent neural activity patterns" (Keppler, 2018b, 3) and that "*the brain produces an individual stream of consciousness by periodically modifying the ZPF*" (Keppler, 2013, 3). It should be pointed out that this self-consistent explanatory approach can be extended beyond conscious perception to incorporate also self-referential consciousness (Keppler, 2018b), altered states of consciousness (Keppler, 2018a,b), as well as declarative memory functions (Keppler, 2020).

CONCLUSION AND OUTLOOK

The strength of the novel cosmopsychist paradigm presented here lies in the bridging of the explanatory gap the conventional materialist doctrine struggles with. This is achieved by proposing a comprehensible causal mechanism for the formation of phenomenal states that is deeply rooted in the foundations of the universe. More specifically, the sort of cosmopsychism we advocate brings a new perspective into play, according to which the structural, functional, and organizational characteristics of the NCC are indicative of the brain's interaction with and modulation of a UFC. In this respect, the key insights from SED suggest that this field can be equated with the ZPF and that the modulation mechanism is identical with the fundamental mechanism underlying quantum systems, resulting in our conclusion that a coherently oscillating neural cell assembly acquires its phenomenal properties by tapping into the universal pool of phenomenal nuances predetermined by the ZPF. This hypothesis is supported by a large body of empirical evidence.

The novel cosmopsychist paradigm elegantly circumvents the hard problem that arises in prevailing materialist approaches because there are "principled reasons to doubt that phenomenal facts are necessitated by purely structural (or functional or organizational) facts" (Shani and Keppler, 2018, 406). The crucial difference is that in our approach "the relevant structural facts . . . are tasked not with the generation of experience *per se* but, rather, with its modulation and restricted expression" (Shani and Keppler, 2018, 406), leading to well-defined distinctive features between conscious and non-conscious systems as well as conscious and unconscious brain processes. In this context, it should be highlighted that the proposed causal mechanism underlying phenomenal states is predicted to be accompanied by a concomitant phenomenon, namely the emission of characteristic photon pulses (Keppler, 2016, 2018b), paving the way for a new research strategy that aims at corroborating the hypotheses formulated in this paper and eventually ends in the systematic "derivation of psychophysical mapping rules between particular qualia and particular sets of phase-locked ZPF modes" (Shani and Keppler, 2018, 407).

Finally, a note is appropriate with regard to the causal interpretation of our UFC account. According to our approach, the UFC has two complementary description levels (hence the earlier reference to a double-aspect perspective), each of which is coherent in itself. From the physical perspective, the dynamic interaction of the UFC with material systems can be consistently described in terms of energy transfer ensuring causal closure and energy conservation, so that the evolution of the UFC is

fully determined by its physical properties. On this reading, the evolving field passes through a series of physical states and the phenomenal qualities associated with each state can be regarded as accompanying features of the physical processes. On the other hand, even though there remains certainly a lot of work to be done, we believe that our approach has the potential to set the stage for a phenomenological interpretation of dynamical processes, meaning that the processes of which we usually think in terms of physical causation may be self-consistently interpreted in conceptually alternative, phenomenal terms. From this point of view, consciousness may be causally efficacious and turn out to be the ultimate intrinsic force underlying the dynamic transformations described by physics, thus laying the foundations for a scientifically informed idealist worldview.

AUTHOR CONTRIBUTIONS

Both authors contributed equally to this work and approved it for publication.

REFERENCES

Baluška, F., and Levin, M. (2016). On having no head: cognition throughout biological systems. *Front. Psychol.* 7:902. doi: 10.3389/fpsyg.2016.00902

Baluška, F., and Reber, A. (2019). Sentience and consciousness in single cells: how the first minds emerged in unicellular species. *BioEssays* 41:e1800229. doi: 10.1002/bies.201800229

Ben-Jacob, E., Shapira, Y., and Tauber, A. I. (2006). Seeking the foundations of cognition in bacteria: from Schrödinger's negative entropy to latent information. *Physica A* 359, 495–524. doi: 10.1016/j.physa.2005.05.096

Boyer, T. H. (1969). Derivation of the blackbody radiation spectrum without quantum assumptions. *Phys. Rev.* 182, 1374–1383. doi: 10.1103/PhysRev.182.1374

Boyer, T. H. (1975). Random electrodynamics: the theory of classical electrodynamics with classical electromagnetic zero-point radiation. *Phys. Rev. D* 11, 790–808. doi: 10.1103/PhysRevD.11.790

Chalmers, D. J. (1995). Facing up to the problem of consciousness. *J. Conscious. Stud.* 2, 200–219. doi: 10.1093/acprof:oso/9780195311105.003.0001

Crick, F., and Koch, C. (1990). Towards a neurobiological theory of consciousness. *Sem. Neurosci.* 2, 263–275.

Crick, F., and Koch, C. (2003). A framework for consciousness. *Nat. Neurosci.* 6, 119–126. doi: 10.1038/nn0203-119

De la Peña, L., and Cetto, A. M. (1994). Quantum phenomena and the zeropoint radiation field. *Found. Phys* 24, 917–948. doi: 10.1007/BF02067655

De la Peña, L., and Cetto, A. M. (1995). Quantum phenomena and the zeropoint radiation field II. *Found. Phys.* 25, 573–604. doi: 10.1007/BF02059007

De la Peña, L., and Cetto, A. M. (1996). *The Quantum Dice: An Introduction to Stochastic Electrodynamics*. Dordrecht: Kluwer Academic Publishers.

De la Peña, L., and Cetto, A. M. (2001). Quantum theory and linear stochastic electrodynamics. *Found. Phys.* 31, 1703–1731. doi: 10.1023/A:1012670800317

De la Peña, L., and Cetto, A. M. (2006). The foundations of linear stochastic electrodynamics. *Found. Phys.* 36, 350–368. doi: 10.1007/s10701-005-9020-1

De la Peña, L., Cetto, A. M., and Valdés-Hernández, A. (2015). *The Emerging Quantum. The Physics Behind Quantum Mechanics*. Cham: Springer International Publishing.

De la Peña, L., Valdés-Hernández, A., and Cetto, A. M. (2009). Quantum mechanics as an emergent property of ergodic systems embedded in the zero-point radiation field. *Found. Phys.* 39, 1240–1272. doi: 10.1007/s10701-009-9348-z

Desmedt, J. E., and Tomberg, C. (1994). Transient phase-locking of 40 Hz electrical oscillations in prefrontal parietal cortex reflects the process of conscious somatic perception. *Neurosci. Lett.* 168, 126–129. doi: 10.1016/0304-3940(94)90432-4

Doesburg, S. M., Green, J. J., McDonald, J. J., and Ward, L. M. (2009). Rhythms of consciousness: binocular rivalry reveals large-scale oscillatory network dynamics mediating visual perception. *PLoS One* 4:e6142. doi: 10.1371/journal.pone.0006142

Engel, A. K., and Singer, W. (2001). Temporal binding and the neural correlates of sensory awareness. *Trends Cogn. Sci.* 5, 16–25. doi: 10.1016/S1364-6613(00)01568-0

Freeman, W. J. (1991). The physiology of perception. *Sci. Am.* 264, 78–85.

Freeman, W. J. (2004). Origin, structure, and role of background EEG activity. Part 1. Analytic Amplitude. *Clin. Neurophysiol.* 115, 2077–2088. doi: 10.1016/j.clinph.2004.02.029

Freeman, W. J. (2005). Origin, structure, and role of background EEG activity. Part 3. Neural frame classification. *Clin. Neurophysiol.* 116, 1118–1129. doi: 10.1016/j.clinph.2004.12.023

Freeman, W. J. (2007). Indirect biological measures of consciousness from field studies of brains as dynamical systems. *Neural Networks* 20, 1021–1031. doi: 10.1016/j.neunet.2007.09.004

Freeman, W. J. (2009). Deep analysis of perception through dynamic structures that emerge in cortical activity from self-regulated noise. *Cogn. Neurodyn.* 3, 105–116. doi: 10.1007/s11571-009-9075-3

Freeman, W. J., and Vitiello, G. (2006). Nonlinear brain dynamics as macroscopic manifestation of underlying many-body field dynamics. *Phys. Life Rev.* 3, 93–118. doi: 10.1016/j.plrev.2006.02.001

Freeman, W. J., and Vitiello, G. (2007). The dissipative quantum model of brain and laboratory observations. *Electr. J. Theor. Phys.* 4, 1–18. doi: 10.1142/9789812779953_0009

Gagliano, P. (2017). The mind of plants: thinking the unthinkable. *Commun. Integr. Biol.* 10:e1288333. doi: 10.1080/19420889.2017.1288333

Gaillard, R., Dehaene, S., Adam, C., Clemenceau, S., Hasboun, D., Baulac, M., et al. (2009). Converging intracranial markers of conscious access. *PLoS Biol.* 7:e1000061. doi: 10.1371/journal.pbio.1000061

Godfrey-Smith, P. (2013). Cephalopods and the evolution of the mind. *Pac. Conserv. Biol.* 19, 4–9. doi: 10.1071/PC130004

Goff, P. (2017). *Consciousness and Fundamental Reality*. Oxford: Oxford University Press.

Hanlon, R. T., and Messenger, J. (1996). *Cephalopod Behaviour*. Cambridge: Cambridge University Press.

Huxley, T. H. (1866). *Lessons in Elementary Physiology*. London: McMillan.

Keppler, J. (2012). A conceptual framework for consciousness based on a deep understanding of matter. *Philos. Study* 2, 689–703. doi: 10.17265/2159-5313/2012.10.001

Keppler, J. (2013). A new perspective on the functioning of the brain and the mechanisms behind conscious processes. *Front. Psychol.* 4:242. doi: 10.3389/fpsyg.2013.00242

Keppler, J. (2016). On the universal mechanism underlying conscious systems and the foundations for a theory of consciousness. *Open J. Phil.* 6, 346–367. doi: 10.4236/ojpp.2016.64034

Keppler, J. (2018a). Shedding light on the fundamental mechanism underlying hypnotic analgesia. *Ann. Palliat. Med.* 7, 170–176. doi: 10.21037/apm.2017.04.03

Keppler, J. (2018b). The role of the brain in conscious processes: a new way of looking at the neural correlates of consciousness. *Front. Psychol.* 9:1346. doi: 10.3389/fpsyg.2018.01346

Keppler, J. (2020). The common basis of memory and consciousness: understanding the brain as a write–read head interacting with an omnipresent background field. *Front. Psychol.* 10:2968. doi: 10.3389/fpsyg.2019.02968

Lyon, P. (2015). The cognitive cell: bacteria behavior reconsidered. *Front. Microbiol.* 6:264. doi: 10.3389/fmicb.2015.00264

Marshall, T. W. (1963). Random electrodynamics. *Proc. R. Soc. London A* 276, 475–491. doi: 10.1098/rspa.1963.0220

Marshall, T. W. (1965). Statistical electrodynamics. *Proc. Camb. Phil. Soc.* 61, 537–546. doi: 10.1017/S0305004100004114

Mathews, F. (2011). "Panpsychism as Paradigm," in *The Mental as Fundamental: New Perspectives on Panpsychism*, ed. M. Blamauer (Heusenstamm: Ontos Verlag), 141–156.

Melloni, L., Molina, C., Pena, M., Torres, D., Singer, W., and Rodriguez, E. (2007). Synchronization of neural activity across cortical areas correlates with conscious perception. *J. Neurosci.* 27, 2858–2865. doi: 10.1523/JNEUROSCI.4623-06.2007

Nagasawa, Y., and Wager, K. (2017). "Panpsychism and priority cosmopsychism," in *Panpsychism: Contemporary Perspectives*, eds G. Brüntrup, and L. Jaskolla (New York, NY: Oxford University Press), 113–129. doi: 10.1093/acprof:oso/9780199359943.003.0005

Nakagaki, T., Yamada, H., and Tóth, Á (2000). Maze-solving by an amoeboid organism. *Nature* 407, 246. doi: 10.1038/35035159

Reid, C. R., Latty, T., Dussutour, A., and Beekman, M. (2012). Slime mold uses an externalized spatial "memory" to navigate in complex environments. *Proc. Natl. Acad. Sci. U.S.A.* 109, 17490–17494. doi: 10.1073/pnas.1215037109

Rodriguez, E., George, N., Lachaux, J. P., Martinerie, J., Renault, B., and Varela, F. J. (1999). Perception's shadow: long distance synchronization of human brain activity. *Nature* 397, 430–433. doi: 10.1038/17120

Seth, A. K., Izhikevich, E., Reeke, G. N., and Edelman, G. M. (2006). Theories and measures of consciousness: an extended framework. *Proc. Natl. Acad. Sci. U.S.A.* 103, 10799–10804. doi: 10.1073/pnas.0604347103

Shani, I. (2015). Cosmopsychism: a holistic approach to the metaphysics of experience. *Philos. Pap.* 44, 389–437. doi: 10.1080/05568641.2015.1106709

Shani, I., and Keppler, J. (2018). Beyond combination: how cosmic consciousness grounds ordinary experience. *J. Am. Philos. Assoc.* 4, 390–410. doi: 10.1017/apa.2018.30

Singer, W. (2015). "The ongoing search for the neuronal correlate of consciousness," in *Open MIND*, eds T. Metzinger, and J. M. Windt (Frankfurt: MIND Group).

Tononi, G., and Koch, C. (2008). The neural correlates of consciousness: an update. *Ann. N. Y. Acad. Sci.* 1124, 239–261. doi: 10.1196/annals.1440.004

Trewavas, A. (2014). *Plant Behaviour and Intelligence*. New York, NY: Oxford University Press.

Tyndall, J. (1879). *Fragments of Science: A Series of Detached Essays, Addresses, and Reviews*. London: Longmans, Green, and Co.

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

METHODS
published: 30 April 2020
doi: 10.3389/fpsyg.2020.00855

The Problem of the Task. Pseudo-Interactivity as an Experimental Paradigm of Phenomenological Psychology

*Alexander Nicolai Wendt**

Department of Psychology, Heidelberg University, Heidelberg, Germany

Real-life problems are almost always socially complex, even when we are by ourselves. Psychological problem-solving research must therefore integrate complexity as a domain of investigation. However, the simulation of complex interactions represents a major challenge to designing experiments dealing with the nature of social interaction: Simulated social interaction, even when enacted by confederates, is not identical to the actual social interaction. Subjects will tend to enact simulated interaction in distinct ways. To understand these differences, the different situation enactments ought to be analyzed psychologically. Essentially, an instruction to perform in an experimental setting cannot guarantee that the experimental subject will take a certain attitude toward the situation. Early psychology of thought considered the social nature of the experimental situation when discussing the notion of the task. Modern experimental psychology can draw on these reflections in order to grasp better the essential characteristics of social complexity and to establish pseudo-interactivity as a phenomenologically enriched experimental paradigm. Its methodological power is illustrated by an exploratory experimentation on problem-solving.

Keywords: phenomenological psychology, problem-solving, semantic complexity, pseudo-interactivity, psychology of thought

OPEN ACCESS

Edited by:
Christopher Gutland, Sun Yat-sen University, China

Reviewed by:
David Alexej Tobinski, University of Duisburg-Essen, Germany
Lars Allolio-Näcke, University of Erlangen–Nuremberg, Germany

***Correspondence:**
Alexander Nicolai Wendt
alexander@puwendt.de

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

Received: *04 January 2020*
Accepted: *06 April 2020*
Published: *30 April 2020*

Citation:
Wendt AN (2020) The Problem of the Task. Pseudo-Interactivity as an Experimental Paradigm of Phenomenological Psychology. Front. Psychol. 11:855. doi: 10.3389/fpsyg.2020.00855

INTRODUCTION

In the last century, the psychology of thought has partly developed into the psychology of problem-solving (for a historical and sociological overview see Kusch, 1999). In the beginning of this development, approaches like *Denkpsychologie* from Würzburg investigated the functions of higher cognition beyond associationism. Scholars, such as Karl Marbe, August Messer and, most importantly, Karl Bühler, proposed psychology of thought to be a field of research that dealt with the specifics of conscious experience in response to tasks: easy or complex. Thus, the notion of "task" became fundamental for their investigations.

The investigation of problem-solving is the externalist heritage of thought-psychology, mainly drawing on the notion of the task as it has been used, for example, by the De Groot (1946) in his seminal work on the psychology of chess, "Het denken van de schaker." In his conceptual reflections, de Groot tends to use the terms of "task" and "problem" interchangeably. It is no wonder that Newell and Simon (1972), who had read de Groot's thesis on chess before writing "Human Problem Solving," also do not make a clear distinction between "task" and "problem."

Their groundbreaking *opus magnum* inspired all the contemporary research on problem-solving, implicitly bridging it with psychology of thought.

Unlike psychology of thought, recent investigations no longer consider non-inductive explanations. The original thought psychology, on the other hand, received its original inspiration by alternative epistemological approaches: Oswald Külpe, the founder of the *Würzburgian* Institute of Psychology and the father of psychology of thought, stated in his historical retrospection on the movement: "In epistemology, it is the problem of reality that came thought-psychology's way. Already before the experimental investigation of thought, it had been remarked, especially by Twardowski, Husserl, Freytag, that the content of thought and its object are different and that it is not directed at itself but at something transcendent beyond its own sphere" (Külpe, 1922, p. 320; translation by the author).

It is a recurrent motif in the initial psychology of thought literature to refer to Husserl (see Münch, 1998), namely in Marbe (1901), Messer (1906), and Bühler (1907), but also in Selz (see Seebohm, 1970) and Lindworsky (1916). Thus, there may be a developmental trajectory of the phenomenological part of thought psychology that is similar to the psychological development via de Groot to Newell and Simon. But this is not only a historical alternative. Problem-solving research seems to have encountered an impasse, if not a deadlock (as it has been described by, e.g., Getzels, 1982; Quesada et al., 2005; Ohlsson, 2012; Funke, 2014; Wendt, 2018). Thus, the question must be posed whether the contemporary limitations of cognitivist problem-solving research result from the neglect of the epistemological problems that were apparent in the days of thought-psychology's foundation.

To give an example, problem-solving research frequently presupposes the general motivational directedness of the experimental subjects toward the instructed situation, the so-called "properly motivated subject" (Newell and Simon, 1972, p. 54). Wertheimer labeled this presupposition as "constructed foundations" (for a discussion see Nerney, 1979). In other words, the researchers simply assume that giving an instruction means that the subjects will have the problem. Put more appropriately (and honestly), problem-solving research is not concerned with whether or not its subjects really have a problem in the authentic sense of the word or simply pass their time in the laboratory. It is a methodological concern if this difference can be neglected since it cannot be measured with precision, i.e., as a behavioral variable. Unlike contemporary cognitivists, the original psychologists of thought have at least discussed this question.

The original conundrum can be summarized as follows: Does it require theoretical and epistemological considerations in order to advance problem-solving research as an instance of behavioral cognitivism in psychology? Without making the case for any, merely ideological, response that advocates either empiricism or rationalism, it should be conceded that the way to an answer must directly deal with the psychological phenomena themselves. This approach demands a minimal adjustment to the inductive methodology of experimental psychology. This step may be called a phenomenological parallax. It does not entail a rigid rationalism, even though most radical empiricists may think so. Its basic contention is simply that the scientific concept formation is not self-given. Therefore, it is necessary to investigate the foundation of scientific concepts by intuition and reflection.

In the case of problem-solving research, this means that the experiential foundations of notions like "task" or even "problem" themselves (see Wendt, 2018) should be restored in order to facilitate both theoretical debate and theory-directed empirical investigation. It is not enough to rely on empiricist paradigms just because they give supposedly reliable or effective results. Their phenomenal relevance should be discussed if not proven beyond empirical traditions. Only this emancipation from any irrelevant experimental approaches will allow psychology to be a rigorous science and to establish criteria of quality that surpass the statistical control of internal consistency.

What Is a Task?

From an externalist point of view, it might seem as though a task and, not differently, an experimental task could be adequately described as a material constellation. For example, Philipp and Koch claim that "the term task can be basically understood as 'what subjects have to do in an experiment.'" (Philipp and Koch, 2010, 383). This common conception presumes a constellation of the experimental subject and a goal state: "the link between a task and a goal is that a task can be assigned by a third party. [. . .] It is then the duty of every single person to decide whether he or she accepts the task assignment. If he or she does, the depersonalized task becomes a personal goal of that specific person" (Künzell et al., 2018, p. 6).

Quite obviously, the notion of the goal is no less ambiguous than the notion of the task. Taken within an externalist framework, it entails both the instruction, i.e., an imperative communication by the examiner, and its representation: "the instructions given to subjects in an experiment must define the task(s) at a level that permits comprehension of what has to be accomplished" (Schneider and Logan, 2014, p. 29). Ultimately, the priority of this external input, this stimulation or stimulus of the instruction, is the onset and condition for any representational information processing, that is, problem-solving (in the respective sense).

Consequently, an externalist framework must attribute all variation in the motivation of subjects to differences in their representations of the same instruction. This explanatory approach is artificial since it overrates the separation between representations of the instruction and motivational processes that do not relate to the instruction. Furthermore, the construction of an instruction may make it empirically impossible that all subjects experience the same goal-directedness.

Most importantly, however, the externalist explanation cannot provide a sufficient understanding of the motivational dynamics of tasks: assuming that a task conveys an instruction via communication to a goal, the question remains why a person would want to allow this conveyance. In fact, this gap could even be understood as a second task, namely, the task to accept the experimental task. This explanatory insufficiency results in an infinite regress because the externalist notion of the task only presupposes the actual transformative faculty of the task. This is something that cannot be grasped from an externalist standpoint.

Psychology of thought, on the other hand, did not overlook the complexity of the conscious experiences and processes that underlie the experimental situation. Three members of the movement made important and successively more complex contributions to the understanding of these particular dynamics: Henry J. Watt in his "Experimental contributions to a theory of thought" from 1904, Narziß Ach's "On the agency of volition and on thought" from 1905, and Otto Selz' "On the laws of the orderly course of thought" from 1913. Bearing major methodological resemblance, the three texts are dedicated to the investigation of thought in its immanent psychological nature. Thus, the notion of the task is not only relevant as an experimental condition but in its psychological function.

Henry J. Watt's Observation of the Task

Watt's approach was the critical evaluation of associationist explanations of thought. Associationist psychologists, like Georg Elias Müller or Theodor Ziehen (see Müller, 1913), had claimed that the responses to a given stimulus word could be explained exclusively by local associative links in memory. These links obey the associative laws in the tradition of Alexander Bain, David Hume and, ultimately, Aristotle. In other words, the traditional claim was that reproductions (sc. associations) result entirely from the strength of associative links that were given by a learning experience or disposition.

In his experiments, Watt revised this assumption under the condition of restricted, instead of free, association. At this point, it is essential to highlight the associationist conviction that there is an innate mechanism to respond to a stimulus. In this sense, the traditional notion of the task was trivial: When one sees a word or any other stimulus, they cannot but produce a reaction. Although it might not be entirely unfounded to assume humans, or even life itself, have a universal responsive nature, this mechanism cannot account for the specificity of experimental tasks because it is necessarily unrestricted, or, as Watt would say, free. This difference is reflected in the distinction between free and restricted associations. Likewise, it reflects the fundamental difference between the associative response mechanisms and tasks in the proper sense.

It becomes clear that, for Watt, a mere stimulus itself does not constitute a task – it is solely a free association[1]. Restricted associations become possible when the stimulus word is accompanied by a task in the specific sense of the word, such as finding "a superordinate concept," "a subordinate concept," "a whole," "a part," etc. Watt, then, observed that the variance of experimental responses almost entirely depended on the restriction given by the task, overriding the unspecific effect of the free association. His results show, for example, that the tasks of finding "a whole" or "a part" invite significantly more imaginative thoughts than the tasks of finding "a superordinate concept" or "a subordinate concept". In contrast, in their introspective reports, the subjects reported more verbal experiences for these latter tasks.

An example can illustrate this observation: Given the stimulus "apple," a subject is more likely to have a verbal association of the word "fruit" given the task "superordinate concept" but given the task "a whole," the imaginative, i.e., visually imaginative, association of a tree becomes more likely. These results might not come as a surprise, but they contradict the traditional contention that the stimulus word's associative links account for the responsive variance. Even Ziehen's broadening of the associationist framework by his theory of associative constellations does not encompass the situational autonomy of the responses created by a particular task.

Watt concludes from his observations that the associationist explanation should be rejected. With it, Watt repudiates the bundle theory of consciousness, claiming that a continuous consciousness is "the condition for the occurrence of more complex factors, the task being one of them" (Watt, 1905, p. 422; translation by the author). Yet, his focal point remained the notion of the task. Apart from his observations about the variance in the experimental responses, he makes an important observation: "As we have seen, before and after the stimulus word with previous preparation, there occurs a pause, either a pause of waiting for the stimulus word or of waiting for the searched or appearing imagination" (Watt, 1905, p. 430; translation by the author). He, then, claims that this pause is the empirical manifestation of the consciousness of the task: "a task, thus, is a state of consciousness that exists in order to determine a sensible series of reproductions, and that can only be indicated as this (series), even comes to consciousness only as it" (Watt, 1905).

Mayer and Orth (1901), as well as Marbe (1901), first described a specific state of mind that was not identical with either volition or imagination:

> "Experimental subject Mayer made the observation of an unspecific conscious process after hearing the stimulus word 'poetic meter' that was followed by the spoken word 'trochee.' In other cases, the subject succeeded to further characterize this experience. Orth observed that the stimulus word 'mustard' evoked such a peculiar conscious process which he meant to characterize as a 'memory of a common figure of speech.' It was followed by the reaction 'seed.' In all these cases, the subjects could not notice the presence of imaginations in the conscious state by which they specified the psychological phenomenon. We subsume all these manifold processes despite their obviously entirely different qualities under the name of 'states of consciousness' [*Bewusstseinslagen*]" (Mayer and Orth, 1901, p. 6; translation by the author).

This particular observation of *Bewusstseinslagen* triggered a fundamental debate about the role of intuition (in the sense of sensational, imaginative content) in thinking. In psychology of thought this crucial debate found its climax in the works of Bühler (1907, 1908a,b). However, the notion of *Bewusstseinslage* is also pertinent for the question of what is the task since Watt claimed that the pause before and after the stimulus can be identified with these "states of consciousness." In other words, the consciousness of the task cannot be identified with a simple

[1]Going farther than Watt, August Messer concluded that the idea of a purely free association is unrealistic: "In the first experiment, the task was given to associate the first arbitrary word that comes to mind. However, it can be observed most frequently that the subjects were searching for a word that stood in a meaningful relation to the stimulus word, even without being aware of it. Thus, they subconsciously set themselves a more specific task" (Messer, 1906, p. 22; translation by the author).

imagination or volition, it is a particular state of mind. A state of mind that is an example of the specific form of experience that turned out to be the main topic for *Würzburgian* psychology of thought. However, just like Marbe, Watt did not succeed in characterizing the phenomenon further. The breakthrough for the understanding of the phenomenon of *Bewusstseinslage* was Bühler's work that drew on Husserlian phenomenology.

Narziß Ach's Conception of the Task

The essential result from Watt's investigations was the rediscovery of the task as a psychological phenomenon and its integration into the understanding of conscious life, implicitly refuting all externalist conceptions. The immanent psychology of the task could bridge the gap between instruction and actual goal-directedness. However, due to the associationist heritage of his idea and investigation Watt could not reach any further. One limitation of his approach was the imposed separation of the task from volitions as distinct states of consciousness. In contrast, Ach explicitly directed his research interest at the investigation of the will. This might seem unexpected for anyone who takes "psychology of thought" to be an exclusively intellectualist endeavor in the antiquated sense of a separation between faculties of the soul. Regardless, the *Würzburgian* psychology of thought never reduced the scope of its research merely to cognitive processes.

Nevertheless, like Watt, Ach's starting point was influenced by associationist methodology. Continuing the work of his teacher Müller from Göttingen, he planned to complement the principles of connections between imaginations rather than raising doubts about the fundamental idea of connections between insulated mental elements. His investigations led him to the assumption of three principles, the first two are identical with Müller's associationism: the associative and the preserving tendency. The third, Ach's discovery, were the determining tendencies: "Determining tendencies can be understood as the effects that result from the particular imaginative content in the imagination of the goal and entail a determination in the sense of the meaning of this content of imagination" (Ach, 1905, p. 187; translation by the author).

The Brentanoesque distinction between an act and its content is the key to the idea of determining tendencies. Ach distinguishes the non-imaginative act of determination from the imaginative contents of consciousness. In order to tackle the prior, he coins the notion "being-conscious" (*Bewußtheit*) that has also been translated as "consciousness of objects" (Meyer, 1924) or "awareness" (Mason, 1913) and should not be confused with consciousness (*Bewußtsein*):

> "With the help of the method of systematic experimental self-observation, we have obtained results for the analysis of the content of consciousness that repeatedly showed that a complex content of knowledge was present simultaneously. Withal, this knowledge was given without intuition, i.e., it did not contain phenomenological aspects, such as visual, acoustic, kinesthetic impressions or memory pictures of such impressions that would qualitatively determine the content that is given as knowledge. Such results occurred for all subjects who attended the systematic experimental self-observation. We shall call this being-present of a non-intuitional knowledge 'being-conscious' [Bewußtheit]" (Ach, 1905, p. 210; translation by the author).

The specific act of "being-conscious" is directed at non-intuitional content and the determining tendencies are "knowledge," which is one of these contents. Despite this act-psychological distinction, Ach's explanation remains teleological and the motif of determination is conceived as a strict mechanism: "being-conscious is the stronger, the greater the level of arousal of the imaginations that are at disposition, the stronger the arousal of the tendencies of reproduction" (Ach, 1905, p. 218; translation by the author). Consequently, Ach conceptualizes the determining tendencies as actual causal determination, ultimately forfeiting the independence from one-by-one associations that had been discovered by Watt – maybe because little to no attention had been paid to the autonomy of the act.

Otto Selz' Investigation of the Task

Ach had stated that "the effect of the determining tendencies does not only originate from present intentions but these tendencies can also be brought about by suggestive influence, commando, or by the task" (Ach, 1905, p. 196; translation by the author). Thereby, continuing Watt's approach of integrating the notion of the task into an overarching unity of experience that does not depend on the internal-external dichotomy. Despite making progress in the analysis of the content of experience, his mechanistic teleology falls short on a phenomenological consideration of complexity on the side of the act. A further step was taken by Selz.

Selz explicitly criticizes both Watt and Ach for not having emancipated themselves sufficiently from the associationist traditions. Nevertheless, his work does not dismiss their line of investigation since he adopts the question: "What are the laws by which the determining tendencies cause the orderly course of the intellectual processes?" (Selz, 1913, p. 3; translation by the author). His answer, however, overcomes some of the limitations of his predecessors. In the center of his attention was the notion of the task, making him, among the *Würzburgian* psychologists of thought, the most elaborate commentator on the topic.

Selz' first progression is empirical. Watt and Ach had investigated the variance of behavior under the condition of different tasks, separating different experimental groups by their respective tasks. Historically speaking, this step was necessary because it is the first occurrence of the phenomenon and, therefore, he could not manipulate it. Selz, however, determined the separation of groups is a weak point since the individual preparation cannot be investigated. Instead, he decided to alternate the tasks with every trial so that "at least the majority of the subjects has to find the solution instead of reproducing a solution that is already prepared" (Selz, 1913, p. 10; translation by the author). Consequently, he could examine cognitive activity when a subject is simulataneously faced with the stimulus and the task.

Watt and Ach had taken the task for a singular element of cognition that, despite its determining tendencies, could be identified with the moment of instruction. In contrast, Selz

saw that the stimulus itself, being the material content of the cognitive act, partakes in the formation of experience and is not exchangeable and not a passive subject of association. His basic example is the verbal association of a name: When somebody is searching for the name of a popular person but cannot recall it immediately, another person might give them a hint by spelling the first letters of the name in question. These letters themselves already imply, for example, that the required response is verbal, a meaningful expression in line with certain rules of orthography or traditions of denotation.

Selz, then, makes a step of abstraction. He claims that these basic qualities are present even when the stimulus is very rudimentary, for example, when the subject only hears incomprehensible mumbling that, nonetheless, indicates the application of language. Thus, he concludes that a scheme of the reaction is present already in the stimulus material. The availability of this scheme, however, does not determine a factual response without the corresponding task. If the stimulus is presented without a task, the conscious state of the subject can be characterized as a "blank form" (see Selz, 1913, p. 218), a metaphor Selz uses to illustrate the nature of the scheme.

Likewise, a task without a stimulus leaves the subject in a comparable, but not identical, state of a "blank form": If there is just a stimulus, the subject is prone to lose attention. If there is just a task, the subject will probably experience a tensed orientation toward the pending stimulus. Selz conceptualizes this orientation, which is induced by the consciousness of the task, as goal-orientation. Only the interaction of the task and the stimulus fulfils the determining tendencies, namely it determines the subject. Selz calls this interaction the "total task" (*Gesamtaufgabe*). The emphasis on the totality reflects the way in which he transgresses the associationist idea of the task as an element in a constellation. Instead of a constellation, Selz speaks of a complex.

A complex is a totality that resembles the idea of gestalt. It is a form that does not reside in any of its parts alone but in their entirety. Thus, a subject who perceives a vacancy within a complex may apprehend the task is to complete the complex. Selz describes this apprehension as the scheme that establishes the field for a response. Therefore, the "total task" is contrived to be the schematic unity of the stimulus and the "task in a broader sense" (with the "broader sense," Selz refers to "the instructions from the examiner the subject must follow"; Selz, 1913, p. 178; translation by the author).

Notwithstanding this, Selz understands the determination of the responsive mechanism as the schematic relation of "complex association." Thus, he does not forfeit associationism on the level of causation. When a stimulus and a task are given in the "total task," the subject must react in a deterministic manner of "knowledge actualization," filling the gaps of the complex. In other words, just as Watt and Ach, Selz remained faithful to the teleological explanations of reproductive thinking in the tradition of associationist psychology. Accordingly, it is safe to say that Selz, despite having critically expanded the notion of the task, did not surprise with his explanation of thought itself. His most valuable psychological legacy is the distinction between the instruction and the "total task," not his mechanistic idea of reproductive and productive thought.

The Task of Problem-Solving

Considering the problem of the task, the crucial question is whether 20th century psychology managed to conserve or even advance the level of reflection reached by the early psychology of thought. Without judging the value of the consecutive work in the field, it must be acknowledged that investigation of thought has undertaken a fundamental transformation, straying from the original discourse. The appeal of cybernetic cognitive sciences overrode the more sedate reflections on the nature of thought. The pioneers of modern-day problem-solving research, Newell and Simon, may have taken Selz into consideration, especially via the lecture of de Groot (Simon, 1999), but they did not reach his conceptual depth (Mack, 1997; van Strien and Faas, 2005). Lacking knowledge about the underlying controversies in the *Würzburgian* psychology of thought, they could not grasp its nuances. Instead, they salvaged the available material for their own interests in problem-solving. Ironically, in the process, they replicated some of the conceptual difficulties that had already been discussed by their predecessors.

Introducing imprecise terms like "task environment," Newell and Simon whitewash the conceptual complexity of the underlying foundation: "The term task environment, as we shall use it, refers to an environment coupled with a goal, problem, or task – the one for which the motivation of the subject is assumed" (Newell and Simon, 1972, p. 55). They retreat to "constructed foundations" to try to bracket the motivational conditions of behavior that cannot be bracketed. They also conflate the notions goal, problem, and task.

This becomes increasingly clear in further passages, e.g., when they assume a "very simple problem situation where subjects can (and occasionally do) represent the task internally in quite different ways" (Newell and Simon, 1972, p. 63). Here, they return to an externalist concept of the task that ignores the critical progress made in psychology of thought bridging the internal-external dichotomy. Additionally, Newell and Simon even consider a task environment in the sense of a "Kantian Ding an sich" (Newell and Simon, 1972, p. 56). A task, thus, can be reduced to a "symbol structure" (Newell and Simon, 1972, p. 78) – returning to the associationist contention of a constellationist nature of thought. Furthermore, externalism is accompanied by representationalism:

> "[W]e have insisted that we can know the objective task – 'out there' – only through its particular representations. There is no neutral way to describing the task environment. As a consequence, task instructions do much more than define the task; they provide, in addition, a specific representation of it that can serve to define an initial problem space, and even parts of an initial problem solving program for the subject" (Newell and Simon, 1972, p. 849).

For Newell and Simon, the representation of this external task environment coincides with the "goal" on the side of the subject

as the internal information processing system[2]. Consequently, "[t]he task environment (plus the intelligence of the problem solver) determines to a large extent the behavior of the problem solver" (Newell and Simon, 1972, p. 788). This assertion is a clear setback compared with the earlier psychologists: Even the idea of determining tendencies had served to diversify the cognitive complexity instead of returning to a linear idea of causation. It should be seen as a mental function that is distinct from either associative or perseverant tendencies. The externalist conceptions of Newell and Simon, on the other hand, do not require more than associationist foundations.

In summary, it can be said that the more recent pioneers of psychological problem-solving research did not continue the line of investigation that was established by psychology of thought. The development of psychological research in the following decades, however, revealed that the progress of cybernetics and information sciences do not support psychological progress. Thus, a return to the conceptual depth of psychology of thought is necessary to advance the field.

Empirically Recovering the Complexity of the Task

The important achievements of psychology of thought lie beyond concept formation. The experimental investigations into thinking established a new format of introspective science, the method of "systematic experimental self-observation." In the current discourse about first-person science of consciousness, this contribution to psychology is often underrated. Still, its relevance is more frequently recognized than the potential of phenomenologically revised problem-solving research (see Wendt, 2017). While introspection might have been the most important methodological topic for psychology of thought in the controversial delimitation from elementarist, functionalist, or behaviorist psychology, the discourse within the approach itself also considered more specific subjects, such as the nature of the task. With regard to present-day problem-solving research these latter considerations are more useful than a mere plea for a return to introspection. To put it another way, problem-solving research can harness the contribution of psychology of thought without the necessity of radical methodological concessions. This is the epistemological background of pseudo-interactivity.

Pseudo-interactivity tries to restore the discourse about the task as it has been undertaken in psychology of thought as a basis of problem-solving research. Its premise is the experimental observation of problem-solving and decision-making by the means of simulation that has been employed for the last decades in laboratory research with computers, especially in the context of so called "complex problem-solving" (Dörner and Funke, 2017). The basic experimental configuration is an imaginative scenario or a game – structurally comparable with the investigations on chess or cryptarithmetic by Newell and Simon. In its current form, this means that subjects are confronted with a digital "micro-world" (Funke, 1993) which simulates a more or less arbitrary content, such as the administration of a city (Dörner et al., 1983) or the scheduling of a daily routine (Holt et al., 2011).

Unlike traditional approaches that focus on algorithmic complexity, pseudo-interactivity reconsiders the meaning of the simulation in the light of the nature of the task itself. Drawing on Selz, the question is what schemes come to the fore when an experimental subject is asked to imagine they were, for example, the manager of a fictitious business. In the debate of the last decades, the difference between certain scenarios was a matter of formal difference, especially regarding the particular problem space. The paradigm of pseudo-interactivity, in contrast, shifts the attention from the possible operations of solving within a certain "micro-world" to the experiential conditions of the situation in which a subject partakes in a problem-solving task. Consequently, it is not decisive whether or not the subjects actually reach a possible solution or even improvement. Rather, pseudo-interactivity is designed to investigate the specifics of the experience, which allow a certain scenario to successfully simulate a problem.

METHOD

Pseudo-interactivity is an experimental paradigm that encompasses behavioral studies, e.g., based on computer simulations. Its main purpose is the investigation of experiential differences between phenomenologically distinct types of situations, such as problems, challenges, and fatalities (see Wendt, 2017). Pseudo-interactivity allows for quantitative and qualitative measurements and it is not restrictive on the specific design. However, unlike comparable paradigms, it requires an explicit conceptual decision concerning the relationship between the material content, especially the task, and the experiential conditions of the experimental subjects.

The architecture of pseudo-interactive experiments combines two fields of psychology. First, it complements the field of complex problem-solving (CPS) research. CPS has worked with digital simulations in order to simulate so called complex problems. Funke uses five qualities to distinguish complex problems from simple problems: intransparency, dynamics, connectivity, polytely, and complexity (Funke, 2012). These qualities can be formalized and, thus, used as structural principles of a simulation. For example, a digital scenario is intransparent if the experimental subject cannot access all operating parameters that underlie the simulation as algorithms. Similarly, a simulation fulfils the quality of dynamics if the parameters, e.g., the arithmetic relation between inputs, change throughout the experiment. Therefore, the complexity of CPS can be effectively characterized as algorithmic complexity.

Algorithmic complexity, or, more specifically, algorithmically simulated complexity, is a term that is meant to describe a certain understanding of situations, such as the situation in a psychology laboratory, as complex systems. Dörner (2008) draws on the analogy of real-world complexity and a spring mattress in order to explain complexity: "everything is tied with everything else and nobody knows exactly how" (Dörner, 2008, 285; translation

[2]This idea, however, does not neglect much of the previous discourse because Watt, Ach, and Selz did not overcome their teleological convictions, either. This lack of critical flexibility reveals an important difference between the rather cognitivist group of *Würzburgian* psychologists of thought and a phenomenological alternative, such as Messer (1908, 1911), Lindworsky (1916), and, later, Graumann (1955, 1960).

by the author). He continues describing complexity by some characteristics, such as "a great many variables", "variables being 'interconnected,' or 'weak' causal relations." What makes this understanding of complexity amenable to algorithmic implementation is the basic contention that complexity depends on the number, variety, and connectivity of "variables." Dörner explicitly promotes that this form of complexity can be simulated by a machine, viz. a computer. The essential property of this kind of simulation is "the mathematical formulation of the hypotheses about the connections between the variables" (Dörner, 1996, p. 505; translation by the author). Algorithmic complexity, thus, is a mathematical representation of real-world inter-relatedness. Present-day CPS paradigms rely on this understanding of complexity as the criterion of validity for its simulations.

However, algorithmic complexity does not indicate the actual experience of complexity. The accumulation of a confusing number of ever-changing parameters does not guarantee that the experimental subject will actually experience a shift of attitude in comparison with, for example, playing chess. Accordingly, algorithmic complexity remains a label for a certain class of simulations that differ materially from so called simple problems, but not necessarily experientially. Certainly, the implicit motivation for the design of these simulations was the search for ecologically valid replications of real-life problems. However, the reality of problems does not (only) derive from their material multifariousness but from their vividness. The attempt to create more accurate and realistic simulations alone will not lead to more authentic experimental behavior.

In order to deal with this predicament, the second architectural principle is phenomenology. Unlike the search for greater ecological validity, phenomenology is not directed at the structural similarity between the simulation and an external situation. Instead, it tries to understand the essence of the experimental situation as a genuine experience in the lifeworld of an experimental subject. Hence, phenomenological psychology reinvigorates that inheritance from the psychology of thought.

The first reflective step of phenomenology has to be the critique of the CPS paradigm. What are the experiences that correspond with the respective moments of a simulation? The crucial insight is that no experimental subject actually encounters an intuitive problem when confronted with the cover story of a simulation, such as the administration of a business. Even in the fictitious case of an entirely ingenuous person who does not consider or question the experimental content, the best case will always be a projective imagination. Asking a person to imagine that they were the mayor of a city, as did Dörner and colleagues, can only result in imagination. This is the case if the instruction worked, that is, if it effectively manipulated the intention of the experimental subject, or if the subject willingly consented to follow the instruction (in the sense of a "hidden dialogue," Lyons, 1970). Ultimately, however, this means that no simulation succeeds in presenting a factual scenario. The only immediacy a subject can experience is the laboratory situation – and therefore the constructed foundations of the simulation themselves. These foundations, however, are the reality of the task because of the nature and prerequisites of any laboratory situation. In other words, all the experimental content of problem-solving research is the result of a communicative influence or agreement to simulate.

Pseudo-interactivity does not try to resolve this constraint, but takes it as its point of departure. Instead of a more potent instruction or cover story, its design starts with the contention that the entire simulation relies on projection. Knowing that a person who engages with a CPS simulation can only imagine the situation they are supposed to be facing by projection, the need to construct an ecologically valid cover story is relieved. Instead, the decisions about the experimental design must concern the nature of the task. More precisely, in line with Selz, the composition of a pseudo-interactive experiment takes into consideration that all the content of the experiment forms the "total task." Therefore, it becomes possible to design "total tasks" that express phenomenologically distinct modes of situations.

In order to manifest these situational modes (Wendt, 2017) in the experimental design, the core element of pseudo-interactive experimentation is semantic complexity instead of algorithmic complexity. While CPS experiments treat their parameters with a certain reluctance in order to cater to the common sense of the experimental subjects, pseudo-interactive experiments allow for extraordinary experimental behavior since it does not conservatively replicate the constraints of a plausible scenario. Only if there is a decisional margin may the experimental behavior represent the experiential diversity of simulated situations. In algorithmic complexity, there is only a margin for eventual solutions, not for the initial attitude.

Semantic complexity (viz. semantically simulated complexity), on the other hand, does not commit to the connectionist contention that complexity emerges from the fuzzy interaction of variables. However, it also does not reject it. Rather, it is a complementary conception of what happens in a complex situation. The fundamental idea that characterizes the semantically complex understanding of complex situations is "sense-making" and, more precisely in this context, "participatory sense-making" (De Jaegher and Di Paolo, 2007): "the process of generating and transforming meaning in the interplay between interacting individuals and the interaction process itself" (Fuchs and De Jaegher, 2009, p. 466). In a communicative situation, this intersubjective sense-making becomes its own realm of complexity– a complexity which can be simulated semantically.

In the context of CPS, the most important aspect of this specific form of complexity is that it cannot be reduced. "Sense" is manifested in meaningful actions that do not consist of elements that could be algorithmically represented as variables. As a consequence, psychological experiments that try to engage with this side of real-life complexity must employ different means. Semantically simulated complexity is one such means and tries to get a hold of "participatory sense-making" via the simulation of the semantic subtlety of communication. Pseudo-interactivity strives to cover this complexity in order to provide access to the nuances of experience.

How, then, can the appropriate experiential complexity be provoked? What tasks invite a variety of experiential attitudes? Pseudo-interactivity employs the simulation of personal interaction. Unlike practical decisions, such as the scheduling of

daily routines, communicative interaction is genuinely manifold, and subjects may express themselves in a fictitious dialogue. Drawing on Fuchs, this form of fiction can be explained as "extended empathy":

> "To begin with, it entails an explicit, cognitive operation, namely, the conscious envisioning of the situation of the other, which often employs information about him that one could not infer directly from the situation at hand. Also, it involves an imaginative operation, that means, a transposition into an 'as-if' scenario (i.e., as if I were the other) which transcends the bodily or physical level. As a result, it seems necessary to differentiate between a primary, implicit, or bodily empathy and an expanded, explicit, or imaginative empathy. The latter already involves a certain degree of virtuality" (Fuchs, 2014, p. 158).

However, the operational advantage of CPS is that the experimental input is quantitative so that it can be used as a parameter for the algorithmic simulation and measured without transformation. However, it is not impossible to maintain these advantages without having to return to algorithmic complexity. One way would be to program a code that accommodates a free communicative input. While increasing the decisional margin, comparability is lost. Another way is to offer an ample variety of pre-coded verbal operators that offer a sufficiently wide margin of decisional alternatives and allows anticipation of the semantic complexity.

The following description and discussion of one of the first pseudo-interactive experiments employs this second design. In the given methodological context of this article, the main purpose of discussing this experiment is to illustrate the capabilities of pseudo-interactivity. The actual relevance of the paradigm for CPS, especially the theoretical background and the hypotheses that were investigated, will not be presented. It would require a lengthy explanation about the phenomenology of the problem (for a general outline see Wendt, 2017) in order to explain the precise meaning of the manipulations and measurements, distracting from the methodological purpose of the present discourse. The corresponding discussion about the experiment can be found elsewhere, alongside further experimentation (see Wendt, 2019).

Exploratory Experimentation

A first pseudo-interactive experiment was conducted in early 2019 with a sample of 40 (34 female, 6 male, age m = 27, 3, s = 11, 4) students from Heidelberg University. Its premise was the investigation of the difference between an urgent and solvable situation ("problem") and an urgent but unsolvable situation ("fatality"). In order to create continuity with CPS simulations, it was based on a variety of well-known experimental settings ("classical problems"), such as the "cannibals and missionaries" game[3] (see **Supplementary Figure S2**), that were transformed to fit into the global setting of the simulation. It was implemented with the coding language MATLAB, version 2018a, and the integrated toolbox "psychtoolbox."

The global setting was a traveling scenario. Participants were asked to imagine that they were traveling in a Spanish speaking country in South America and were to meet their friend at the train station within the next half an hour. Throughout the experiment, they could navigate the representation of their position on a map of the city Maracaibo in Venezuela (see **Supplementary Figure S1**). On their course, they would encounter up to four of the "classical problems" mentioned above. However, these "classical problems" were not presented by an instruction but as an encounter with a simulated person. The (pseudo-) interaction with this person could be executed by the application of 100 pre-coded operations, such as "concentrate oneself" or "provoke somebody".

Most importantly, no action was demanded by an instructive task, neither in the beginning of the experiment nor in the case of an encounter within the simulation. Also, no action was required to finish the experiment. After 30 min, the simulation was terminated automatically, and the participants were asked to answer conclusive questions about their experience. Since the experimental design was inherently open, a variety of inputs were measured, such as the course of the participants on the maps and the actions in the "classical problems." The most important measure, however, was the selection from the 100 operators (for a detailed discussion of the method see Wendt, 2019).

The difference between the two conditions was that the "problem" condition could solve the "classical problems" while the "fatality" condition could not. This difference was implemented by respective communicative responses by the simulated persons who accompanied the "classical problems." For example, an elderly man who represented the "classical problem" of the "tower of Hanoi" asked the participant in a written message to help him carry a fragile machine. In the "problem" condition, it was possible to solve the task in the same fashion as one would solve the "tower of Hanoi." In the "fatality" condition, the elderly man would interrupt the process after some actions and return all the parts of the machine back to their initial location. Hence, the difference did not exist on the logical level, i.e., the algorithmic architecture of variables, but on the operational and communicative level, i.e., the semantic material that is available for the participants' sense-making.

In both conditions, the participants could find identical solutions to the "classical problems" by themselves. The "fatality" group of participants, however, could not implement their solution into the simulation. Based on phenomenological considerations about the problem (see Wendt, 2019), it was expected to find preferences for certain operators depending on the experimental group. Instead of discussing the hypotheses in detail, it is of greater importance at this point to revisit the methodological meaning of pseudo-interactivity.

While most research on problem-solving behavior presupposes that the task can be understood as a definite element of the experimental process, pseudo-interactivity investigates the development of the "total task" in a simulation with a wide decisional margin. Whether or not some simulation will be experienced as a problem in the emphatic sense of

[3]Logic transformations in the setting of three missionaries who must escort three cannibals over a river. They may use a boat for two persons. It is prohibited that there are more cannibals than missionaries on either of the sides of the river at any point in time.

the word, cannot be guaranteed by instructions. Nonetheless, the interactions with the simulations are expressions of the sense making that occurs throughout experimentation and they express the particular attitude of the participants, the mode of their situation. Thus, it is necessary to include indicators of these attitudes into experimentation on CPS. Moreover, the method allows one to investigate the situational modes of experience themselves.

Taken together, the pseudo-interactive experiment was designed to make the original sense making of subjects who are confronted with a simulation accessible for psychological research, i.e., observable and measurable, instead of assuming "constructed foundations." A step-by-step comparison of the pseudo-interactive and a typical CPS design may help to highlight the critical differences:

Introductory Phase

The purpose of the introduction in CPS used to be the unambiguous instruction. Briefing the participants is important for comparability. In contrast, in pseudo-interactivity the Selzian contention that the "total task" develops in confrontation with the stimulus material is taken into consideration. Therefore, the introduction does not give an unambiguous instruction although it remains clear that any content will already evoke a "blanket form" for eventual behavior. However, the purpose is to maintain the process of instruction incomplete.

The General Situation

The validity of CPS research depends entirely on the compliance of the participants. If a participant succeeds to solve a problem by chance or with, for example, a playful attitude, the results will be misguiding. Pseudo-interactivity does not depend on compliance but describes it in a phenomenologically refined form. Because of the decisional margin, the creativity of the participants can burgeon. If a participant disregards the introductory narrative and wants to test the experimental coding by the execution of mischievous patterns of behavior, they may do so, and it will reflect in their data. The purpose of the entire examination is to investigate under what circumstances certain forms of "total task" will be experienced.

Problem-Solving Phase

The research on CPS concentrates on the manipulation of parameters that can be judged as either favorable or unfavorable. Consequently, an optimization algorithm can standardize any CPS simulation. The operators in pseudo-interactivity, on the other hand, are not subject to a metric norm. If a person chooses certain communication actions to progress over others, they cannot be optimized in the mathematical sense of the word. However, from the point of view of coding, it is not a mistake to underpin the semantic complexity with an algorithmic architecture so that the responses of the simulation are strictly deterministic and thereby guarantee diagnostic objectivity and statistical comparability. The entire meaning of the actions results from the "pseudo-interaction," i.e., the imaginative and projective cognitions of the participants.

Measurements

Complex problem-solving research normally investigates the input of integers that can be used as metrics for the statistical interpretation. Semantic complexity in its present operationalization does not have an obvious statistical measurement apart from the number of selected operators. Yet, given a theory of situational behavior, generalizations are possible. In the given experiment, a phenomenological theory of the problem, called "structure of problematic situations" (see Wendt, 2019), was used to evaluate the selection of operators. It includes five dimensions that distinguish different modes of situations, such as problems, challenges, and fatalities. These dimensions were "serious vs. playful," "burdensome vs. comfortable," "exploratory vs. committed," "objective vs. subjective," and "active vs. passive." An independent rating of the 100 operators that were used in the experiment resulted in a distribution of representativity for each operator on the five dimensions. As a consequence, each selection of an operator in the simulation could be interpreted as an approximative expression on the respective situational dimension, if it had been rated accordingly.

Data-Model

The underlying data-model is the result of the two-step validation of the 100 operators (see above under "Measurements"). In the first step, the five behavioral dimensions that had been derived from phenomenological reflections, were used to sort the 100 operators. In the second step, this sorting was validated by expert rating. As a result, the operators that deviated by 1 to 2 standard deviations from the average score of the dimension, were given a score of 1, and the operators that deviated by more than 2 standard deviations from the average score were given a score of 2. This generalization of the ratings was meant to compensate for outliers in the ratings, increasing reliability. Consequently, the selection of each operator during the experiment can be measured on the respective dimensions with a score from −2 to 2 units (the meaning and scale of this unit derives from the standard deviation of the expert ratings). The resulting scores can be interpreted in isolation as a single event or in partial or complete aggregation. For example, a participant might have a cumulative score of 5 on the dimension "exploratory vs. committed" over the course of the entire experiment, meaning that she or he selected more operators that were rated as exploratory than those that were rated as committed.

Introspective Reports

A legacy of psychology of thought is that modern problem-solving research has employed the controversial method of "think-aloud-protocols." The purpose of the data is to determine the detailed process structure of problem-solving, such as typical reactions to difficulties. The introspective data from pseudo-interactive experimentation, however, do not serve an equally specific purpose. Instead of the path of solution, the research is directed at the process of goal setting or problem finding. Hence, the pseudo-interactive manipulation serves the purpose of finding discrete modes of situations, such as distinct forms of

goal setting and problem finding; a proviso is that the differences between the experimental conditions are as small as possible.

RESULTS

Due to the methodological focus of the investigation, reporting the results is primarily supposed to illustrate the conceptual significance of pseudo-interactivity. An interpretation of these tentative results as a contribution to the concept formation of problem-solving and problem finding has been published elsewhere (Wendt, 2019).

Of primary interest is a descriptive account of the situational difference between the participants that were faced with a "fatality" and those who could solve the "classical problems" that they were encountering. All 40 participants voluntarily used 1581 operators (there was no obligation to use them). Curiously, the distribution of operators between the two experimental groups was almost even (788 in the "problem" group and 793 in the "fatality" group). Overall, some operators enjoyed greater popularity than others, as can be seen in **Figure 1**. The most frequent operators were "ask somebody for help" (82 times), "question somebody" (75), and "understand the circumstances" (73). The least frequent operators were "strictly judge somebody" (1), "deny something" (2), "provoke somebody" (2).

These results indicate that the selection was not entirely random but guided by semantics. This is not trivial since the algorithmic function of the operators did not differ. In other words, on the side of the simulation, there was no difference between "asking" or "provoking" somebody. Yet, from the participants' point of view, it was not easy to understand that the operators were redundant since the underlying mechanism was opaque and the reactions of the program did not repeat.

The differences between the two experimental groups can be described by the specific operator usage that did vary between them. For example, the group faced with "fatality" preferred the operators "evade the situation" (5 to 2), "limit oneself" (6 to 1), or "avoid something" (6 to 2). However, due to the low base frequency, these differences do not bear great explanatory weight. The opposite case shows a clearer pattern. The group of participants who could solve the "classical problems" showed prominent preference for operators, such as "reflect all circumstances" (17 to 8), "investigate something" (12 to 6), "question somebody" (43 to 32), "research into something" (16 to 7), or "take initiative" (25 to 16).

Whilst not being the only operators with ostensible differences between the groups, these operators are of special interest because they belong to the group of operators that were considered salient to the dimension of "exploratory vs. Committed." Accordingly, the overall change in the pattern of this dimension, which reflects operator preferences that relate to exploratory attitudes and behavior, helps to understand the situational difference between the two experimental groups. The general change can be illustrated by a line chart that represents the change over the course of the 30-min experiment (**Figure 2**).

The diagram shows the development of the average score on the dimension "exploratory vs. committed" by groups over the course of the 30-min experiment as a compound score of the rating for all selected operators. Generally, the preference of both groups tends toward operators that have been rated exploratory rather than committed. Yet, the participants of the "fatality" condition show a less pronounced tendency toward this extreme. On average, the operators which they applied during the entire experiment had an average score of about 2 units while the participants of the "problem" condition had a cumulated score of 5 units (for the exact procedure of acquisition for the unit of measurement see "Data-model" above).

The diagram shows that there is almost no difference for the operator preferences in the first minutes. This is consistent with the manipulation because the first difference in the encounters with the simulated persons who present the "classical problems"

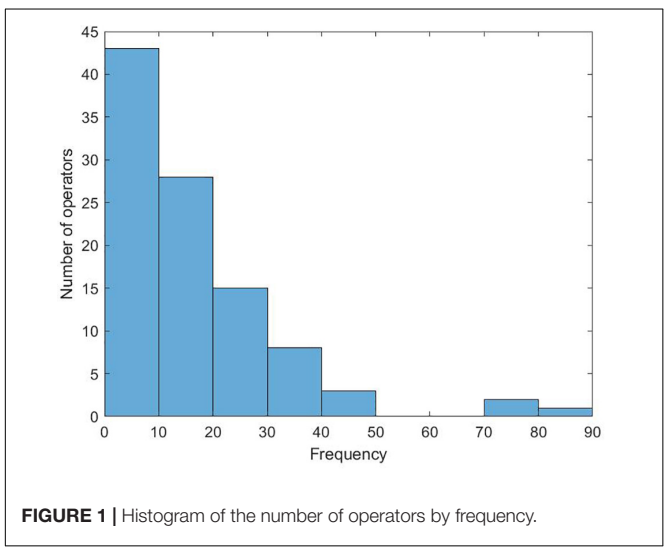


FIGURE 1 | Histogram of the number of operators by frequency.

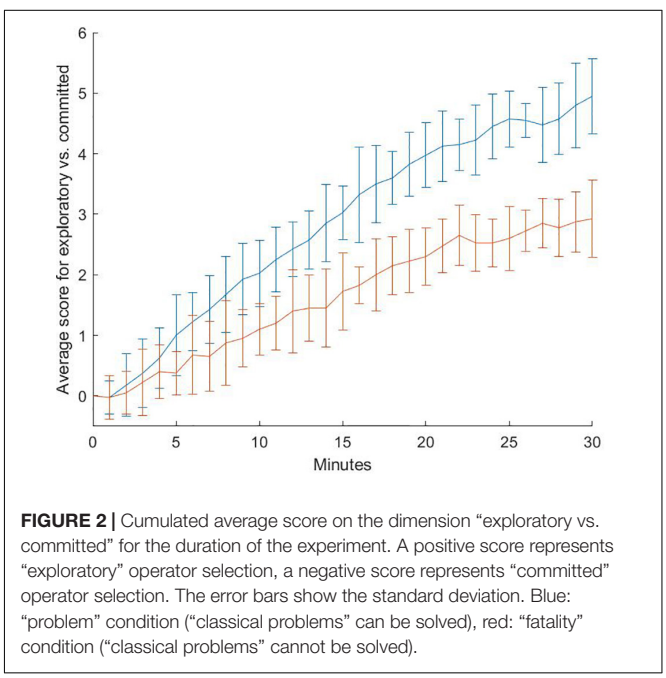


FIGURE 2 | Cumulated average score on the dimension "exploratory vs. committed" for the duration of the experiment. A positive score represents "exploratory" operator selection, a negative score represents "committed" operator selection. The error bars show the standard deviation. Blue: "problem" condition ("classical problems" can be solved), red: "fatality" condition ("classical problems" cannot be solved).

happens after 240s. From this point on, the two groups show a continuously growing gap of their average "exploration vs. commitment" score. For the total number of operators that were selected by the two experimental groups, there is a significant difference in the present sample on this dimension, $t(38) = 1.93$, $p < 0.05$ [medium effect size $d_{Cohen} = 0.51$ (0.19–0.83)]. However, this score is a composite of operator choices that score either on "exploration" or on "commitment." A closer look can give a more fine-grained resolution of the decisional patterns (**Figure 3**).

These diagrams show that the difference between the two groups cannot be reduced to the participants in the "fatality" group either choosing less "exploratory" operators or choosing more "committed" operators. Actually, the composite score reflects a tendency for both extremes of the dimension. Consequently, the "fatality" condition can be described as an experimental situation in which participants have greater preference for committed operations, i.e., restricted and fixed actions, and lesser preference for exploratory operations, i.e., actions of discovery and experimentation, than participants of the "problem" condition[4].

These quantitative aspects of the investigation are reflected in the introspective reports[5]. Participants in the "fatality" condition characteristically reported frustration and aggravation when asked about their general impression: "I tried to help and not achieving success was frustrating" (18 years old male); "I thought that I could help but everyone was unfriendly, so I kept moving" (23 years old female); "In some situations I felt desolated" (19 years old female). When asked about the experimental situation, several responses expressed reactance or even reluctance: "It lacked concrete instructions" (21 years old female); "There was no direction to the train station" (26 years old female); "I would have liked a better image of the city. In reality, one would see the surroundings and not a GPS" (20 years old female).

Participants in the "problem" conditions gave different responses. Despite having completed a simulation that was almost identical to the other experimental group, their reports reflect motivation and immersion: "I wanted to prove that I could move without fear in a foreign city" (22 years old female); "I had a good feeling after helping the merchant" (23 years old female); "I had the intention to help, so I felt useful" (23 years old female). Comments on the experimental situation did not express dissatisfaction but constructive criticism: "I would have liked to see some more things that were happening in the surrounding area" (49 years old female); "I liked that I actually had influence, for example, when calming the persons. I was not interested in music or the noise of the streets. Being able to choose between good and bad actions, was a good feature" (25 years old female), "I would have liked to play more and make more moves" (same person).

DISCUSSION

Clearly, differences between experimental groups concerning the apprehension of the general situation are a secondary effect that is inherent to all experimental designs. However, there are few approaches for the systematic investigation into the immanent structure of these differences and its experimental manipulation. Hence, exploratory results of pseudo-interactive research cannot be validated easily by empirical comparison. Some theoretical approaches have been provided by the discourse about the person-situation-dichotomy, for example, Pawlik (1978), or Lantermann (1980), or Frederiksen (1972). Yet, these postulations do not offer a method of validation but only models for the interaction between persons and

[4]These results are in accordance with the hypotheses posed by the structure of problematic situations (Wendt, 2019). These hypotheses and their rationalization do not have to be discussed on occasion of these primarily methodological considerations.

[5]All translations by the author.

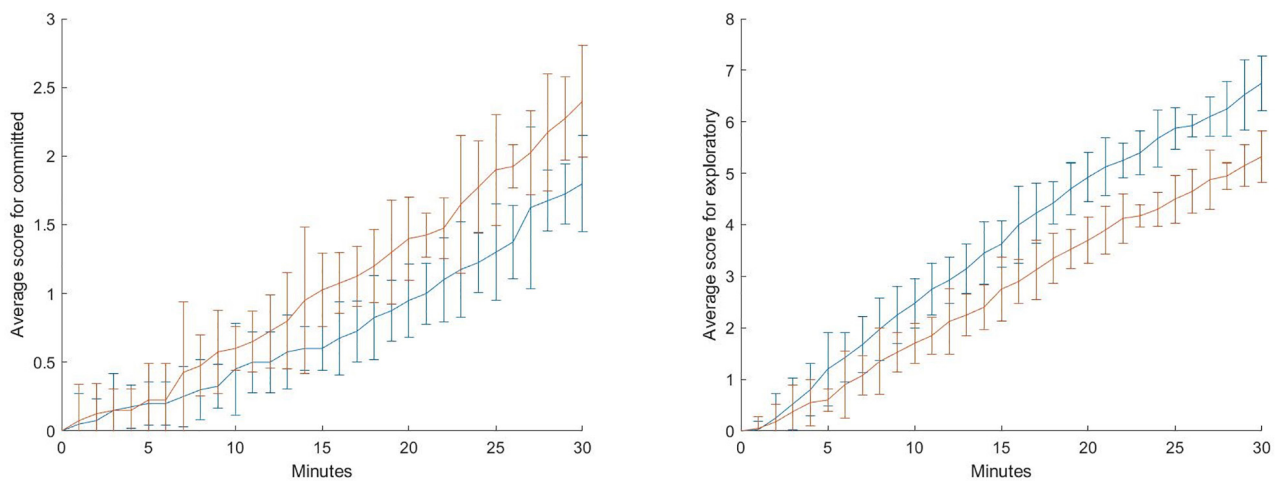


FIGURE 3 | Cumulated average score for "committed" **(left)** and "exploratory" **(right)** for the duration of the experiment. The error bars show the standard deviation. Blue: "problem" condition ("classical problems" can be solved), red: "fatality" condition ("classical problems" cannot be solved).

objectified situations. A more pertinent form of validation is phenomenological psychology.

Unlike empirical concepts of the situation, such as the DIAMONDS-model (Rauthmann et al., 2014), phenomenology does not rely on the, by necessity, restricted range of empirical data. Rather, phenomenological psychology draws on eidetic reflection to grasp the nature of its subject matter. The contributions on the topic of the situation by, for example, the Utrecht school of phenomenology (Buytendijk, 1954; van den Berg, 1955; Linschoten, 1963), show that situatedness is an essential property of experience. Therefore, the empirical results made by applying pseudo-interactivity as an experimental paradigm cannot be seen as mere contingencies. On the contrary, they deal with a structure of experience that necessary for any further investigations of, for example, problem-solving. Only if the experiential characteristics of a certain experimental setup are examined, can interpretations about the subjective attitude toward the experimental content be justified.

In other words, research that relies on common sense in the construction and validation of their experimental designs bear the risk of uncontrolled mistakes about the actual situation they are creating. Pseudo-interactive investigations are a way to cope with this risk without having to rely on purely reflective considerations about the nature of the situation. The decisive step is to introduce a situational alternative that allows a comparison. Consequently, pseudo-interactivity allows psychology to move beyond the "constructed foundations" of experimental sciences. The peculiarities of the general situation in a laboratory, which can be abstractly described by notions like "compliance," "demand characteristics," or "social desirability," can be described and compared in a concrete and phenomenologically adequate fashion (for a detailed discussion see Wendt, 2018).

The value of these descriptions and comparisons, however, depends on a return to the psychology of the task brought about by the *Würzburgian* psychology of thought. Watt's fundamental insight was that tasks are not external to consciousness. If modern problem-solving research wants to be faithful to his conclusions, no formal criterion for experimental designs can be established that would guarantee that the experimental subjects conceive the laboratory situation as a task. The obvious backdoor of psychological interpretation is to assume that all recorded data conforms to the salient behavior. Yet, this division of behavior based on measurement must be arbitrary. Moreover, it constrains the psychological observation to predicted reactions. It disregards the creative responsibility of science.

Despite having touched on the phenomenon, Watt did not systematically investigate its relation to cognition. This step was taken by Ach. He claimed that the consciousness of a task could be explained as a determining tendency. For present-day psychology, this means that it is not enough to assume a single relative principle for all cognitive functions. Unlike tendencies of association, determining tendencies are characterized by anticipation. Thus, the emergence of task-consciousness should not be mistaken for a linear causation. Rather, it requires self-referential relations and thus subjectivity in the phenomenological sense of the word (for an understanding of circular causation see Fuchs, 2017).

Ach's explanation still maintained an analogy between the associative, perseverant and determining tendencies. It was Selz who emphasized the border between associationist psychology and psychology of thought by claiming that the task was not a simple state of consciousness but a whole that should not be understood as a constellation but as a complex. Subsequently, the actual cognitive mechanism that may explain determining tendencies is a schematic actualization of knowledge. Problem-solving research might learn from this step that the situation created in an experimental setup cannot be predictably modified by changing singular elements. The experience that an experimental subject will have when confronted with a laboratory situation will necessarily be complex and difficult to predict. Thus, a rigorous empirical approach to investigate these complex subjective dynamics is required since common sense assumptions are neither reliable nor controllable. Pseudo-interactivity helps bridge the conceptual gap. It tries to make the subjective experiences of the experimental subjects traceable, or, to borrow a term from ethnomethodology, "accountable."

The presented exploratory experiment that tries to distinguish the experience of a "fatal" from a "problematic" situation, demonstrates a certain resemblance between pseudo-interactivity and ethnomethodology. Ethnomethodology tries to discover the exact process of creating rules that give structure to everyday life. These rules and norms are considered "methods" from the subject's experiential point of view: "the most important assumption that drives ethnomethodological approaches is the methodic and orderly character of everyday activities that appear chaotic and messy at first glance" (Reeves et al., 2016, p. 330). However, ethnomethodology is restrained to a sociological perspective and shies from introspective reports and considerations about psychological processes.

Another like-minded project is the particular ecological psychology that has emerged from phenomenological psychology (e.g., Graumann and Kruse, 1998). It discusses the notion of the situation in a holistic fashion, but it is emancipated from the rather individualistic thought psychology in order to embrace social psychology. Still, its contributions can help to understand better the complexity of the situation as a meaningful and complex aspect of life. Likewise, anthropological psychology (e.g., von Uslar, 1973) describes the existential conditions of situatedness. Nevertheless, none of these approaches has developed an experimental nexus to current problem-solving research that carries the conceptual heritage of psychology of thought. Pseudo-interactivity may fill this gap. However, the experimental designs have to be improved and validated.

The present investigation demonstrated the utility of pseudo-interactivity. It foreshadowed the major challenges for the paradigm: the experimental design requires a degree of precision and structure that could not yet be reached. For example, the success rate of the "classical problems" is at 17.9%. Thus, the fatality by design due to practically unsolvable encounters might be conflated with apparent insolubleness by difficulty. The future development of comparable experiments will shed light on the practicability and the scope of the paradigm.

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. The patients/participants provided their written informed consent to participate in this study.

AUTHOR CONTRIBUTIONS

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

FUNDING

The author acknowledges the financial support of the Deutsche Forschungsgemeinschaft and Ruprecht-Karls-Universität Heidelberg within the funding program Open Access Publishing.

ACKNOWLEDGMENTS

The author is indebted to Dr. George Hausmann for his valuable linguistic contributions.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fpsyg.2020.00855/full#supplementary-material

REFERENCES

Ach, N. (1905). *Über die Willenstätigkeit und das Denken*. Göttingen: Vandenhoeck & Ruprecht.

Bühler, K. (1907). Tatsachen und Probleme zu einer psychologie der denkvorgänge: über Gedanken. *Arch. Gesamte Psychol.* 9, 297–365.

Bühler, K. (1908a). Über Gedankenzusammenhänge. Archiv für die gesamte Psychologie, 1–23.

Bühler, K. (1908b). Über Gedankenerinnerungen. Archiv für die gesamte Psychologie, 24–92.

Buytendijk, F. J. J. (1954). Avant-Propos. *Situation* 1, 7–14.

De Groot, A. (1946). *Het Denken van de Schaker*. Amsterdam: Noord-hollandsche Uttgevers Maatschapij.

De Jaegher, H., and Di Paolo, E. (2007). Participatory sense-making: an enactive approach to social cognition. *Phenomenol. Cogn. Sci.* 6, 485–507. doi: 10.1007/s11097-007-9076-9

Dörner, D. (1996). "Der Umgang mit unbestimmtheit und komplexität und der Gebrauch von Computersimulationen," in *Umweltsoziologie*, eds A. Diekmann, and C. C. Jaeger, (Opladen: Westdeutscher Verlag), 489–515. doi: 10.1007/978-3-322-99626-8_21

Dörner, D. (2008). "Umgang mit komplexität," in *Industrial Ecology: Erfolgreiche Wege zu Nachhaltigen Industriellen Systemen*, ed. A. von Gleich, (Berlin: Springer), 284–302.

Dörner, D., and Funke, J. (2017). Complex problem solving: what it is and what it is not. *Front. Psychol.* 8:1153. doi: 10.3389/fpsyg.2017.01153

Dörner, D., Kreuzig, H. W., Reither, F., and Stäudel, T. (eds) (1983). *Lohausen: Vom Umgang mit Unbestimmtheit und Komplexität*. Bern: Huber.

Frederiksen, F. (1972). Towards a taxonomy of situations. *Am. Psychol.* 27, 114–123. doi: 10.1037/h0032705

Fuchs, T. (2014). The virtual other: empathy in the age of virtuality. *J. Conscious. Stud.* 21, 152–173.

Fuchs, T. (2017). *Ecology of the Brain*. Oxford: Oxford University Press.

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

Funke, J. (1993). "Microworlds based on linear equation systems: a new approach to complex problem solving and experimental results," in *The cognitive psychology of knowledge: The German Wissenspsychologie project*, eds G. Fischer, and D. Laming, (Amsterdam: Elsevier), 313–330. doi: 10.1016/s0166-4115(08)62663-1

Funke, J. (2012). "Complex problem solving," in *Encyclopedia of the Sciences of Learning*, ed. N. M. Seel, (Heidelberg: Springer), 682–685.

Funke, J. (2014). "Problem solving: what are the important questions?," in *Proceedings of the 36th Annual Conference of the Cognitive Science Society*, Vol., eds P. Bello, M. Guarini, M. McShane, and B. Scassellati, Austin, TX: Cognitive Science Society, 493–498.

Getzels, J. W. (1982). "The problem of the problem," in *New Directions for Methodology of Social and Behavioral Science: Question Framing and Response Consistency*, eds M. Robin, and Hogarth, San Francisco,CA: Jossey-Bass, 37–49.

Graumann, C. F. (1955). *Die Kriterien des Einfallserlebens*. Köln: Dissertation.

Graumann, C. F. (1960). *Grundlagen Einer Phänomenologie und Psychologie der Perspektivität*. Den Haag: De Gruyter.

Graumann, C. F., and Kruse, L. (1998). "Children's environments: The phenomenological approach," in *Children, Cities, and Psychological Theories. Developing Relationships*, eds D. Görlitz, H. J. Harloff, G. Mey, and J. Valsiner, (Berlin: de Gruyter), 357–369.

Holt, D. V., Rodewald, K., Rentrop, M., Funke, J., Weisbrod, M., and Kaiser, S. (2011). The Plan-a-day approach to measuring planning ability in patients with schizophrenia. *J. Int. Neuropsychol. Soci.* 17, 327–335. doi: 10.1017/s1355617710001712

Külpe, O. (1922). *Vorlesungen über Psychologie*. Leipzig: Verlag von S. Hirzel.

Künzell, S., Broeker, L., Dignath, D., Ewolds, H., Raab, M., and Thomaschke, R. (2018). What is a task? An ideomotor perspective. *Psychol. Res.* 82, 4–11. doi: 10.1007/s00426-017-0942-y

Kusch, M. (1999). *Psychological Knowledge*. New York, NY: Routledge.

Lantermann, E. D. (1980). *Interaktionen. Person, Situation und Handlung*. München: Urban & Schwarzenberg.

Lindworsky, J. (1916). *Das Schlussfolgernde Denken. Experimentellpsychologische Untersuchungen*. Freiburg im Breisgau: Herder.

Linschoten, J. (1963). Die unumgänglichkeit der phänomenologie. *Jahr. Psychol. Psych.* 10, 177–185.

Lyons, J. (1970). The hidden dialogue in experimental research. *J. Phenomenol. Psychol.* 1, 19–29. doi: 10.1186/1752-0509-8-13

Mack, W. (1997). Otto selz und die kognitionswissenschaft. *Brentano Studien* 7, 315–333.

Marbe, K. (1901). *Experimentell-Psychologische Untersuchungen über das Urteil*. Leipzig: Verlag von Wilhelm Engelmann.

Mason, M. P. (1913). Review of allgemeine psychologie nach kritischer methode. *Psychol. Bull.* 10, 285–287. doi: 10.1037/h0064323

Mayer, A., and Orth, J. (1901). Zur qualitativen untersuchung der association. *Z. Psychol. Physiol. Sinnesorgane* 26, 1–13.

Messer, A. (1906). Experimentell-psychologische untersuchungen über das Denken. *Arch. Gesamte Psycho.* 8, 1–224.

Messer, A. (1908). *Empfindung und Denken*. Leipzig: Verlag von Quelle und Meyer.

Messer, A. (1911). Husserls phänomenologie in ihrem verhältnis zur psychologie. *Arch. Gesamte Psychol.* 22, 117–129.

Meyer, M. F. (1924). Review of die form der bewusstheit. *Psychol. Bull.* 21:655. doi: 10.1037/h0067536

Müller, G. E. (1913). Zur analyse der gedächtnistätigkeit und des vorstellungsverlaufs. part 3. *Z. Psychol.* 8, 102–120.

Münch, D. (1998). "Die vielfältigen beziehungen zwischen philosophie und psychologie. das verhältnis edmund husserls zur würzburger schule in philosophie-, psychologie- und institutionengeschichtlicher perspektive," in *Psychologiegeschichte – Beziehungen zu Philosophie und Grenzgebieten*, eds J. Jahnke, J. Fahrenberg, R. Stegie, and E. Bauer, (München/Wien: Profil Verlag), 319–345.

Nerney, G. (1979). The gestalt of problem-solving: an interpretation of max wertheimer's "productive thinking". *J. Phenomenol. Psychol.* 10, 56–80. doi: 10.1163/156916279x00059

Newell, A., and Simon, H. A. (1972). *Human Problem Solving*. Englewood Cliffs, NJ: Prentice-Hall.

Ohlsson, S. (2012). The problems with problem solving: reflections on the rise, current status, and possible future of a cognitive research paradigm. *J. Probl. Solving* 5, 101–128.

Pawlik, K. (1978). "Umwelt und persönlichkeit. zum verhältnis von ökologischer und differentieller psychologie," in *Ökologische Perspektiven in der Psychologie*, ed. C. F. Graumann, (Bern: Huber), 112–134.

Philipp, A., and Koch, I. (2010). The integration of task-set components into cognitive task representations. *Psychol. Belgica* 50, 383–411. doi: 10.5334/pb-50-3-4-383

Quesada, J., Kintsch, W., and Gomez, E. (2005). Complex problem-solving: a field in search of a definition? *Theor. Issues Ergon. Sci.* 6, 5–33. doi: 10.1080/14639220512331311553

Rauthmann, J. F., Gallardo-Pujol, D., Guillaume, E. M., Todd, E., Nave, C. S., Sherman, R. A., et al. (2014). The situational eight DIAMONDS: a taxonomy of major dimensions of situation characteristics. *J. Personal. Soc. Psychol.* 107, 677–718. doi: 10.1037/a0037250

Reeves, S., Greiffenhagen, C., and Laurier, E. (2016). Video gaming as practical accomplishment: ethnomethodology. *Conversat. Anal.Play Top. Cogn. Sci.* 9, 308–342. doi: 10.1111/tops.12234

Schneider, D. W., and Logan, G. D. (2014). "Tasks, task sets, and the mapping between them," in *Task Switching and Cognitive Control*, eds J. A. Grange, and G. Houghton, (Oxford: Oxford University Press), 27–44. doi: 10.1093/acprof:osobl/9780199921959.003.0002

Seebohm, H. (1970). *Otto Selz. Ein Beitrag zur Geschichte der Psychologie*. Heidelberg: Dissertation.

Selz, O. (1913). *Über die Gesetze des geordneten Denkverlaufs*. Stuttgart: Verlag von W. Spemann.

Simon, H. A. (1999). *Karl Duncker and Cognitive science. From Past to Future*. Pittsburgh, PA: Carnegie Mellon University 1, 1–11.

van den Berg, J. H. (1955). *The Phenomenological Approach to Psychiatry. An Introduction to Recent Phenomenological Psychopathology*. Springfield, MO: Thomas O. J.

van Strien, P., and Faas, E. (2005). "How otto selz became a forerunner of the cognitive revolution," in *The Life Cycle of Psychological Ideas*, eds T. Dalton, and R. Evans, (New York: Springer Science), 175–202.

von Uslar, D. (1973). "Ontologische voraussetzungen der psychologie," in *Psychologische Anthropologie*, eds H.-G. Gadamer, and P. Vogler, (München: DTV), 386–413.

Watt, H. J. (1905). Experimentelle beiträge zu einer theorie des denkens. *Arch. Gesamte Psychol.* 4, 289–436.

Wendt, A. N. (2017). On the benefit of a phenomenological revision of problem solving. *J. Phenomenol. Psychol.* 48, 240–258. doi: 10.1163/15691624-12341330

Wendt, A. N. (2018). Is there a problem in the laboratory? *Front. Psychol.* 9:2443. doi: 10.3389/fpsyg.2018.02443

Wendt, A. N. (2019). *Phänomenologie des Problems. Zugleich das Programm für die Verjüngung der phänomenologischen Psychologie*. Heidelberg: Dissertation.

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

CONCEPTUAL ANALYSIS
published: 24 June 2020
doi: 10.3389/fpsyg.2020.01144

Ten Testable Properties of Consciousness

*Christopher W. Tyler**

California Institute of Integral Studies, San Francisco, CA, United States

This article develops a view of consciousness in the context of a new philosophical approach that invokes the concept of emergence, through which the operative principles of each level of organization of physical energy flow are functionally dissociated from those of the levels below it, despite the continuity of the physical laws that govern them. The particular form of emergence that is the focus of the present analysis is the emergence of conscious mental processing from neural activity carried by the underlying biochemical principles of brain organization. Within this framework, a process model of consciousness is developed to account for many of the experienced aspects of consciousness, many that are rarely considered in the philosophical discourse. Each of these aspects is rigorously specified in terms of its definable properties. It is then analyzed in terms of specific empirical tests that can be used to determine its neural substrate and relevant data that implement such tests. The article concludes with an analysis of the evolutionary function of consciousness, and a critique of the Integrated Information Theory approach to defining its properties.

Keywords: consciousness, brain, neural substrate, hard problem, properties, emergence, empirical evidence

Edited by:
Christopher Gutland,
Sun Yat-sen University, China

Reviewed by:
Michael Mascolo,
Merrimack College, United States
Jan Scheffel,
KTH Royal Institute of Technology, Sweden

***Correspondence:**
Christopher W. Tyler
cwtyler2020@gmail.com

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

Received: *18 December 2019*
Accepted: *04 May 2020*
Published: *24 June 2020*

Citation:
Tyler CW (2020) Ten Testable Properties of Consciousness. Front. Psychol. 11:1144. doi: 10.3389/fpsyg.2020.01144

INTRODUCTION

Philosophical Background: Principles of Functional Emergence

Before addressing the properties of consciousness, it needs to be placed in the context of the overall physical reality from which it emerges. This is conceptualized in the form of Emergent Aspect Dualism (Tyler, 2015, 2018, 2019), which reconciles the epistemic dichotomies of monism and dualism, energy and matter, emergence and continuity, neural activity and consciousness, free will and determinism, and even continuous reality with the superposition and multiple worlds interpretations of quantum physics. This philosophical approach takes the view that complex levels of organization of physical energy are both ontologically and functionally emergent from more basic levels (physical energy being defined as the flow of some physical substance, or stored propensity to flow in the case of potential energy[1]). Thus, the forms of energy at play at the level of subatomic physics are kinetic and potential energy, as captured in the continuous Schrödinger Equation. Indeed, the best-known equation in physics is the Einstein Energy Equation, $E = mc^2$, which expresses unity of energy and matter and, in stellar evolution, the emergence of matter from the raw energy of the Big Bang. It is emphasized that the initial references should be consulted for detail of this contextual overview.

[1]It is recognized that, although the Schrödinger energy equation is typically considered the basis of all physics, the only definition of energy in physics is the anemic "work done," which specifies only the magnitude of the energy. Here, the paradigm concept of an energetic process is an energetic particle traveling through the universe, which only does work when absorbed.

As is well recognized, the emergence principle gives rise to the organizational hierarchy adumbrated as consisting of quantum physics, particle physics, macro physics, physical chemistry, biochemistry, cellular biology, neurobiology, neurophysiology, systems biology, and the psychology and philosophy of consciousness. At the level of biology, the organization of cells such as neurons is emergent from the continuity of their component proteins, one major example being the enclosing property of a cell membrane supporting the maintenance of life that is lost when the integrity of the membrane is punctured. The support function of the cells thus emergent from the enclosure of the membrane, a physical instantiation of the Gestalt principle of closure that provides the critical life-sustaining property of a segregated internal environment. This concept of emergence transcends the strong/weak distinction that has become embedded in the philosophy of emergence (e.g., Hartmann et al., 2019; Turkheimer et al., 2019), since it is "weak" in the sense that it is built up step-by-step from its elementary constituents, but "strong" in the sense that an entirely new principle of operation emerges once closure is achieved (Tyler, 2018).

The most immediate form of emergence specific to the human organism is the emergence of consciousness[2], which is the main focus of the present analysis. Specifically, the evidence from our general experience of human mortality, and from neurosurgery in particular, supports the concept that consciousness is an emergent property of the physical activity of the neurons of the brain. In the general conception, this activity is carried by neurotransmitters such as glutamate at the input to neurons and by sodium/potassium ion exchange along their output connections, so these are the physical substrate of the neural activity in question (which ultimately is itself an emergent form of energy flow of the underlying subatomic processes). It is not intended to imply that consciousness is a form of energy *per se*, but a particular form of *organization* of the energy flow of such neural activity.

The substrate of consciousness is assumed to be the neural activity of the brain, but what makes consciousness unique is that it is the only process that "we" know from the internal perspective of what it is like to ***be*** that process (a different form of internal exclusivity). It is this emergent *internal* perspective that entails the hard dualism of the "Hard Problem" of consciousness (Chalmers, 1997), since we cannot take the internal perspective of anything beyond our own brain process (and our own brain process is the specific one in which our consciousness is obligatorily embedded as our internal viewpoint on sensory and working memory information). Thus it is the privacy or exclusivity of our brain process to our own personal subjectivity that entails a dualism that is emergent from fundamentally monistic complex energy function (Hasker, 1999; Tyler, 2018), as opposed to the classic Cartesian dualism of the separate realms of mental and physical substances.

The final dual aspect of this philosophy is the unity of free will and determinism. The primary assumption of the philosophy is that brain function is fundamentally deterministic (though the product of such a complex system that much of the activity is effectively indeterminate noise). However, as MacKay (1960) has shown, even if an all-knowing external system (or "God") had access to all the information in your[3] brain to predict the next optional decision, it could not convey the information to your subjective decision-making capability, or "will," in a form that could necessarily enforce the decision. You would always have the freedom to decide not to follow the all-knowing prediction once it was presented to you. Thus, in your subjective experience, you always have to deliberately make each decision, or deliberately decide to leave the decision to someone or something else. Your free will is inherent, and cannot be removed by an external predictor, no matter how well-informed it is, even though the entire process is, by assumption, fully determinate at the physical level.

MacKay's paradigm resolves these aspects of free will by illustrating that there is no external prediction that necessarily holds force over the internal decision – the process is always subject to a further decision (although the prediction may provide a helpful weighting of the pre-decision factors). And since life involves a continuous series of decisions made by the most complex organ known to nature, it is hard to imagine that the inherent noise variations throughout the sequence of decisions could be considered deterministic throughout a human life.

This analysis might raise the issue of how this paradigm would apply to a non-conscious automaton programmed to make decisions based on accumulated evidence, which could be programmed to process an external prediction as part of its decision inputs (and respond in a way that is not previously predictable). Stating it in this form makes it clear that the issue of free will is separate and independent of the presence of consciousness *per se*, and highlights the question of what the core issue of free will actually is. Is it the question of whether your entire life is laid out (as part of "God's plan"), without you having the power to affect it? Or whether you could in principle access a source that can inform you of the outcome of each decision you have to make in life, to avoid the challenge of having to struggle through the decision process? Although we humans are only recently developing automata with these kinds of capabilities, the same issues could be formulated for such non-conscious systems, although they only seem to be meaningful when viewed from the *internal* perspective.

This brief overview thus outlines how the Emergent Aspect Dualism philosophy simultaneously reconciles monism with dualism, energy with matter, continuous reality with superposition/multiple worlds, emergence with

[2]The concept of consciousness as an emergent function of brain activity is widely accepted in neuroscience and has a long philosophical history, as summarized in the Stanford Encyclopedia of Philosophy https://plato.stanford.edu/archives/sum2015/entries/properties-emergent/.

[3]The "you" in this context should be read as a stand-in for any conscious decision maker.

continuity, neural activity with consciousness, and free will with determinism.

The Role of Consciousness in Quantum Physics

Returning to the physical underpinnings of neural activity, quantal events are treated in standard quantum theory as being in a probabilistic superposition of physical states in which multiple outcomes coexist until an observation made (Feynman, 1985). Emergent Aspect Dualism, on the other hand, treats all these potential outcomes as existing solely in the mental space of the observer (and in the communications of the physics community by whom the probabilities are calculated), but only *one* outcome as having occurred in the underlying reality (Minev et al., 2019). Thus, the unconventional position of Emergent Aspect Dualism is that the many worlds of the resultant outcomes, and their probabilistic superposition, exist only in the mind (or its computational extensions), not in the quantum reality as generally understood[4].

In this way, the emergent dualism that is the outcome of a sufficiently complex process such as the human (or other) brain lies at the heart of the paradoxes of probabilistic quantum physics that supposedly give rise to it. However, applying this philosophical framework to the role of consciousness in quantum physics leads to the conclusion that probability is not a concept that is inherent in physical processes, but an *analytic concept of a human mind* with the memory to accumulate repeated instances of physical events. It is an inherent property of probability that, by definition, it incorporates multiple defined outcomes (of **p** and **not p**, for example) and associates each one with a weight (necessarily based on past experience with those outcomes). These multiple, or complementary, outcomes are therefore in a state of **conceptual superposition** within the specification of probability, *per se*. Since probability, as a mental concept, inherently embeds the superposition of the complementary states, it follows that the Schrödinger Cat paradox and the collapse of the wavefunction are resolved by realizing that the superposition is a property of the mental representation rather than of the physical reality (Schrödinger, 1935; Tyler, 2015, 2019). It should be noted that this clarification of the relationship between consciousness and the properties of the quantum realm is included as an antidote to the widely disseminated concept that the properties of consciousness could derive from the putative non-classical properties of the quantum realm (Wigner, 1970; Penrose, 1995; Bohm, 2002). Nevertheless, the following analysis of the properties of consciousness in terms of classical biochemical processes does not depend in the resolution of that controversy.

A Definition of Consciousness

It is well in developing an analysis of a phenomenon to attempt a definition of the subject matter under investigation. The present treatment is focused on what Block (1995) terms "phenomenal consciousness," the direct experience of being vividly aware of the flow of events (as contrasted with "access consciousness," which corresponds to the information content of mental operations controlling behavior, as in the Integrated Information Theory of Tononi, 2008). Searle (1990) provides an operational definition of phenomenal consciousness as follows; "By consciousness I simply mean those subjective states of awareness or sentience that begin when one wakes in the morning and continue throughout the period that one is awake until one falls into a dreamless sleep, into a coma, or dies or is otherwise, as they say, unconscious." I would extend the wake/sleep distinction for the consciousness definition here to incorporate the distinction of "working memory" (Baddeley and Hitch, 1974), or operational thought, in that consciousness is "what it is like" (Nagel, 1974) to imagine or think about sensory or memory contents at a given moment (as contrasted with all the possible things in memory that we could be thinking about but are presently out of awareness). This contrast marks a major distinction between to the direct operations of phenomenal thought itself, as opposed to the neural organization that is available to contribute to thought (similar to Block's "access consciousness"). All definitions of consciousness are ultimately either ostensive or tautological, but it is hoped that these descriptions help to define the matter at hand relative to the reader's own experience.

Nevertheless, Block's (1995) concept of access consciousness apparently allows that it can in principle exist without incurring phenomenal consciousness. In this sense, it would constitute a form of ***un***conscious information processing that is not distinguishable from what could occur in a (biological or computational) neural network. Indeed, Block's access consciousness thus reads as very close to the operational concept of "working memory," the set of conceptual processes that control speech and behavior. As such, they can be investigated empirically without reference to the conscious experience of the individual under investigation. The present treatment, on the other hand, is focused on the basic set of phenomenally experienced properties of consciousness (which are necessarily those of the author as a consequence of the privacy restriction, but laid out in a form that it is hoped will resonate with the experience of the reader). The basic properties under consideration in the following are defined purely phenomenologically, without reference to the neural properties of the brain at any level of investigation.

ANALYSIS

The Nature of Consciousness

What, then, is the nature of the process of emergent consciousness (C*) that is so characteristic of human experience? In overview, this article will focus on the following ten properties of phenomenal consciousness, providing specific examples of empirically definable tests for the **neural substrate for conscious**

[4]It may be noted that a non-quantized approach to the wave function in the form of the Einstein-Maxwell-Dirac classical field theory is under active development (e.g., Finster et al., 1999; Vegt, 2019).

processing (NSCP[5]), including some classic and some less-recognized properties of consciousness. The concept of the NSCP is distinguishable from the long-established one of the NCC (the neural correlate of consciousness; Crick and Koch, 1990) in that many forms of brain processing can *correlate* with the properties of consciousness without necessarily forming its true neural substrate. It should be noted that the defining property of phenomenal consciousness, its phenomenality (or qualia, in the plural), is not included in the list because it is not clear how it could be testable. The ten NSCP properties are:

1. **Privacy.** The obvious NSCP basis of the privacy of individual experience is its derivation from the separate brain of each individual experiencing C*.
2. **Unity.** Having unified or correlated activity: if a particular set of brain structures is identified as the NSCP, then they should show unitary activity when C* is reportable.
3. **Interrogacy.** One aspect of brain function that has not been investigated is ability to formulate questions, or interrogacy, which seems unique to a conscious mind.
4. **Extinguishability.** The NSCP must exhibit the same time course of complete extinction as does C* itself every time we fall asleep or are anesthetized, and be rekindled when we awake.
5. **Iterativity.** Any plausible NSCP measure must exhibit the iterative cycling through similar states of conscious experience over the experiential range of time scales.
6. **Operationality.** The operationality of working memory is a functional property of the NSCP that is readily accessible to techniques such as behavioral assessment and brain imaging.
7. **Multifacetedness.** Though unitary in its dynamics, the NSCP should exhibit the multifacetedness of the conscious qualia of the sensory field that is characteristic of C*.
8. **Complex interconnectivity.** To be explanatory the NSCP should match the variety of multilevel interconnectivity of conscious experience.
9. **Autosuppressivity.** The attentional suppression that keeps C* moving on from each identifiable mental state to the next is a further property that can be identified in candidate mechanisms for the NSCP.
10. **Self-referentiability.** Human C* has the capability of representing itself within itself, so its substrate has to be able to exhibit the corresponding capability.

These ten properties may be explicated as follows:

1. **Privacy:** One of the irreducible properties of C* is its **privacy**. *Pace* science fiction, as far as we know, there is no way to share our individual C* with anyone else. Verbal and non-verbal forms of communication provide effective means of generating the illusion of sharing C*, but (as too many lovers have found to their cost!), this is only a superficial level of apparent sharing, not a direct experience of another's true internal experience. To meet this criterion, the NSCP must be brain-compatible and must not allow for direct interbrain communication. In the context of quantal theories of C*, this means that the NSCP must not be based on any non-local quantal effects. (Those who accept the non-locality of some superordinate cosmic C* will, however, draw the opposite conclusion).

   Empirically, the privacy of C* is what for a long period was assumed to prohibit meaningful approaches to testing its properties. In recent years, however, the consistency of reports over time and across individuals has been taken to provide sufficient support for its meaningful investigation.[6] For any group of individuals who agree that their internal experience does manifest one of the properties specified here, that property is testable, while ascertaining the proportion that do not agree is an empirical specification of the prevalence that property in the population.

   A standard objection to the communicability of a private experience is Wittgenstein's (1953) argument that a private language would be incommunicable (§256–§271). Here, on the other hand, it is argued that this issue is largely addressed by the ostensive nature of language learning – we develop common concrete concepts by pointing at external examples of them in the external world that are mutually available to the senses across individuals, and build up more abstract concepts by analogy from the concrete examples. The concept of consciousness, likewise, can be communicated by ostensive reference to the difference in mental experience between being awake and being asleep (without dreaming). (Likewise, dreaming *per se* is another private experience, but no-one questions the linguistic communicability of the concept of dreaming).
2. **Unity:** Under ordinary conditions, C* is experienced as **unitary** at a given moment. We have one experience at a time, although we may be able to rapidly switch among multiple experiences over short time intervals. The NSCP must, accordingly, occur either in a single brain site or in a unified neural net of some kind in the brain, rather than in multiple independent brain sites. [Note that neurological split-brain cases are a counterexample that require special treatment based on whether their conscious experience is, in fact, continuously dual or has some other organization (see Gazzaniga, 1985), but this medically instigated controversy will not be addressed here].
3. **Interrogacy:** Though not widely recognized, a defining property of C* is the ability to generate **questions** and represent potential answers. Complex systems other than the brain, such as galaxies, biological organs and the Internet, incorporate extensive recursive interactions and consist of energy processes that undergo development and evolution comparable to those in the brain. Although these

[5]Note that, since the original publication of the distinction of the neural substrate for conscious processing (NSCP; Tyler, 2015) from the neural correlate of consciousness (NCC; Crick and Koch, 1990; Koch, 2004), Tononi et al. (2016) changed their notation from the longstanding NCC terminology to the "physical substrate of consciousness (PSC)" (without providing any justification for the change).

[6]It is assumed that such reports are obtained under fully non-coercive conditions, such that they may be taken on faith to be valid.

systems can be said to process information, however, they cannot meaningfully be said to ask questions. It seems to be a unique property of a conscious system to formulate questions, and a function that gets switched on in humans at about the age of a year. This capability also entails (though perhaps not until a later age) the ability to envisage possible answers in an indeterminate superposition of their probabilistic states of likelihood.

4. **Extinguishability:** A primary property of C* is its lack of continuity. As emphasized by Searle (1990; see "Introduction"), it is **extinguished** every time we fall asleep, are anesthetized or a knocked out by a physical trauma, and is rekindled when we awake (and also, in a somewhat restricted form, when we dream). Although these states are deeply subjective, they can be attested in the form of memory markers of external events (such as our last memory of a radio program before dropping off to sleep). Dreaming is well established as being objectively indexed by rapid eye movements while asleep (REM sleep) and also being extinguished as we awake.[7]
5. **Iterativity:** Another well-established property of C* is its tendency to **iterate repeatedly** through similar states, both when there are problems to be solved (such as anxiously reiterating a worrying scenario) and as a form of pleasure (as in music, which reflects the consciousness of its listeners, and has continual repeats of phrases, themes and whole pieces over a wide range of time scales). C* is thus not a state *per se*, but an iterative sequence of repeated sub-processes, each often entailing a resonance with previous ones.
6. **Operationality:** The operational property is captured by the term "working" in the cognitive science concept of "**working memory**." In other words, "memory" is the ability to store representations of aspects of the world as stable brain states, while "working" is the functionality of not only bringing them to C* but using them to answer questions either in relation to a single memory (first-order) or in relation to the relationships between memories (second-order). The remarkable property of such memories is that we are not conscious of the millions of memories that are maintained in a non-conscious state most of the time. We only become conscious of them when they are accessed by C* for a brief time, which roughly corresponds to the Theater of C* of Baars (1983). There seem to be two forms of access, one being first-order *inoperative* or factual access that is usually included in the functional usage of working memory without involving any operational changes to the stored information (e.g., "what country is Stockholm in?"), and the other being the second-order *operative* access that is well-described as "working memory," to perform some operation on the stored information (e.g., "does a rotated q become a b or a d?").
7. **Multifacetedness:** C* by its nature incorporates all **varieties of human experience**, from logical thought processes and imaginary journey planning through the irreducible qualia of direct sensory and indirect imagery experiences to the array of emotional experiences and primary internal states of C* such as pain and orgasm. Although we still may not be able to envisage what it would mean for the NSCP to exhibit, or possess, such experiences, it is a core requirement of the theory that it would be able to do so. At least in the case of thought or journey planning, the NSCP should be able to exhibit the activation of the specific memory states representing the sequential stages of the specific thought or journey in question.
8. **Complex interconnectivity:** C* is experienced as complexly interconnected, in the sense that each instantaneous state can proceed along many "lateral thinking" paths from any one state to many others (see **Figure 1**). Thus, while the concept of multifacetedness refers to the array of experiential states of C*, interconnectivity refers to the transitional probability among and between these states. This **interconnected flexibility** is part of its generative or creative power. It is not like a finite state machine, that typical proceeds sequentially from any one state to a definite following state. C* is capable of exhibiting multiple connectivity from any facet to many other facets of human thoughts and feelings, unconstrained by logic. (Of course, in some cases well-trodden paths of thought do become established such that C* does operate analogously to a logical finite-state machine, but this may be more the exception than the rule). This property corresponds to the "global workspace" concept of Baars (1983).
9. **Autosuppressivity:** One of the sources of the variety and creativity of C* is that it tends to exhibit the property of burning out at any one state, **suppressing** the tendency to return to that state, thus impelling continuous movement to novelty. This is a well-known property of attention across the visual field ("inhibition of return" Posner and Cohen, 1984; Müller and Kleinschmidt, 2007), and is also a rule of a good writing style, to avoid using the same term or phrase repeatedly in a text. Indeed, this is the opposite of the behavior of a classical finite-state machine, which repeatedly follows the same path from any given state. Autosuppressivity is thus a major contributor to the creativity of humans and other organisms, though it may be overridden by the iterativity property, the tendency to stay in the comfort zone of the same sequential paths of behavior.
10. **Self-referentiality:** A final property of C* is its ability to **represent itself** as a component of the conscious field. This property harks back to Russell's Paradox as a seemingly impossible feat: what is the set of entities that includes itself as a member? But this is a common experience, that we can be (acutely!) aware of ourself

[7] Note that such rapid eye movements are a good example of an NCC, but do not constitute an NRCP since no one would consider such movements of the eyes to be a plausible substrate of consciousness.

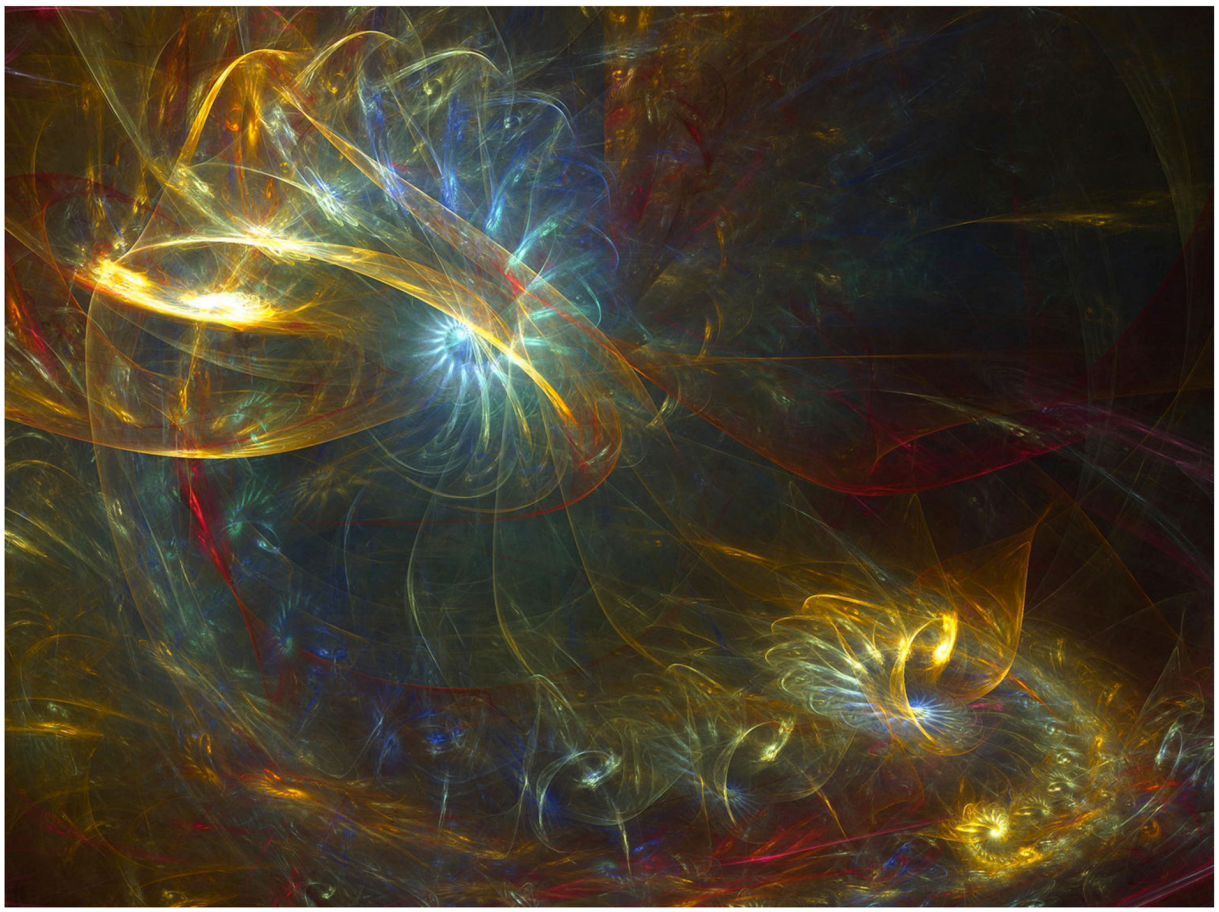


FIGURE 1 | A complexly interconnected iterative and yet unitary structure that represents a dimensional implementation that captures the conceptual nature of the flow of consciousness (C*). Note that the coloration can be viewed as a stand-in for the qualia of the various contents of consciousness. (Free download of colorful-light-swirls-21710.jpg).

as a participant in the field of C*. This property goes beyond the primary quality of the external referentiality of C*, that it has the inherent quality of referring to some form of object outside itself (or what philosophers misleadingly term "intentionality"). C* is experienced as the continuous journey of an identified self, or ego, through the succession of states of experience; that is, not simply an undifferentiated stream of consciousness, but a series of actions and experiences from the viewpoint of an internal entity identified as "me."

EMPIRICAL TESTING FOR THE NSCP

A plausible underlying assumption of the NSCP is that it must have a spatiotemporal isomorphism with the experiential properties of C*. Thus, a core goal in specifying the above properties of C* is to define their spatiotemporal morphology in a testable form. Specific examples of empirically definable tests for the actual NSCP of the properties of C* specified above are as follows:

1. **Privacy.** The obvious basis of the privacy of individual experience is the separation of the brains of each individual. While this test is passed for the typical human brain configuration, it is not easy to set up the converse case, of co-extensive brains for non-private experience. An aspect that relates to this issue, however, is the correspondence between brain states for *comparable* experiences across individuals. When people judge that they have similar individual experiences in particular situations, these similar experiences should be expected to have similar NSCPs in terms of the recordable patterns of neural activity. An experiment along these lines was conducted at the Chinese Normal University of Beijing, in which a group of interacting individuals had their brain activity recorded simultaneously by functional Near-Infrared Spectroscopy (fNIRS) recording (Duan et al., 2015). It could thus be determined if their mutual brain activation patterns were more similar than when recorded asynchronously during non-shared experiences. The non-privacy of the individual brain processes would be validated if there was a relationship among the brain signals corresponding

to their mutually experienced thoughts, when sensory communication was eliminated. Conversely, it is difficult to prove privacy, as it would require a null result from all possible forms of non-sensory communication. One approach is to review the history of studies of extra-sensory perception (ESP), which have all been shown to be a result of manipulation or fraud, despite the best intentions of many of the experimenters (Charpak et al., 2004). Even if a small amount of transmission could be validated under some rare circumstances, historically it has always been found to be such a minute proportion that it confirms the essential privacy of conscious experience under most circumstances.

Empirically, a huge number of studies in human neuroscience are now identifying aspects of brain function in relation to a vast array of stimulation and endogenous environmental conditions. Many such studies, as exemplified by the "mind-reading" study of Huth et al. (2017), relate the brain activity to the reported contents of the individuals' conscious awareness of a defined set of images. Although the accuracy, or information transfer rate, of these studies is low, they do represent a level of external access to the contents of consciousness, suggesting that sufficiently advanced technical system could breach the supposed privacy of consciousness to read minds. However, all such techniques rely on the veracity of the participants' report of their conscious experiences, so their privacy is metained in principle if they choose not to cooperate in such investigations.

2. **Unity.** Consciousness is generally reported to be unitary at any given instant of time. For a particular set of brain structures to be identified as the NSCP, they should either be structurally unitary (such as the anatomical neural net of the claustrum) or have demonstrably correlated (i.e., unitary) activity across the multiple anatomical structures when C* is reportable. To pass this test, the correlation across structures should account for all, or a large proportion of, their recordable activation above the noise level of the recording technique, not just a weak correlation. If the NSCP has a unified *neural substrate*, that substrate should meet the criterion of showing uniform activity as the activity representing the different types of processing fluctuates elsewhere in the brain. If, on the other hand, the NSCP is represented by a particular form of *neural activity* (such as gamma-band energy), that form of activity should be manifested in each of the individual brain areas (cortical or subcortical) at the times identifiable as when the corresponding processing is occurring.

   A key issue arises in terms of the contents of C*, which may switch rapidly over different topics over short periods of time, as in the example of a mother looking out for her children while cooking a meal and mentally preparing for a meeting with an upcoming legal client, with the TV news in the background. If a given individual reports that their C* is literally non-unitary, the corresponding non-unitary switches in activation of the putative NSCP should be identifiable by temporal correlation techniques.

3. **Interrogacy.** One aspect of brain function that does not seem to have been investigated either philosophically or empirically is the process of formulating questions, or interrogacy. Coming up with questions is a creative process that all philosophers and scientists engage in professionally, yet it does not seem to have been codified as a psychological process or studied in a neuroimaging context. Although question-generation is an established field of study in the educational field (Davey and McBride, 1986; Rosenshine et al., 1996), it has yet to become a topic of investigation the domain of cognitive neuroscience.

   The first requirement is to develop a protocol for putting an individual in a controlled state of question generation. Participants would be asked to think of a question about some topic that they have not previously formulated, and indicate when they have come up with a completed formulation. The panoply of brain imaging techniques can then be brought to bear on the issue of the particular substrate of the question-generation component of C*, based on the time period immediately preceding the question-generation completion time. The NSCP should be coextensive with the brain processes underlying the interrogacy activity, once it is studied.

4. **Extinguishability.** The NSCP must exhibit the same time course of complete extinction as C* itself every time we fall asleep or are anesthetized, and be rekindled when we awake. This association could be tested with a button monitor that has to be held down while we are falling asleep or being anesthetized, but will be released by the muscular relaxation with the onset of the sleep state. In humans, this would be most easily tested with continuous scalp electroencephalography (EEG) recording but could be attempted with functional Magnetic Resonance Imaging (fMRI).

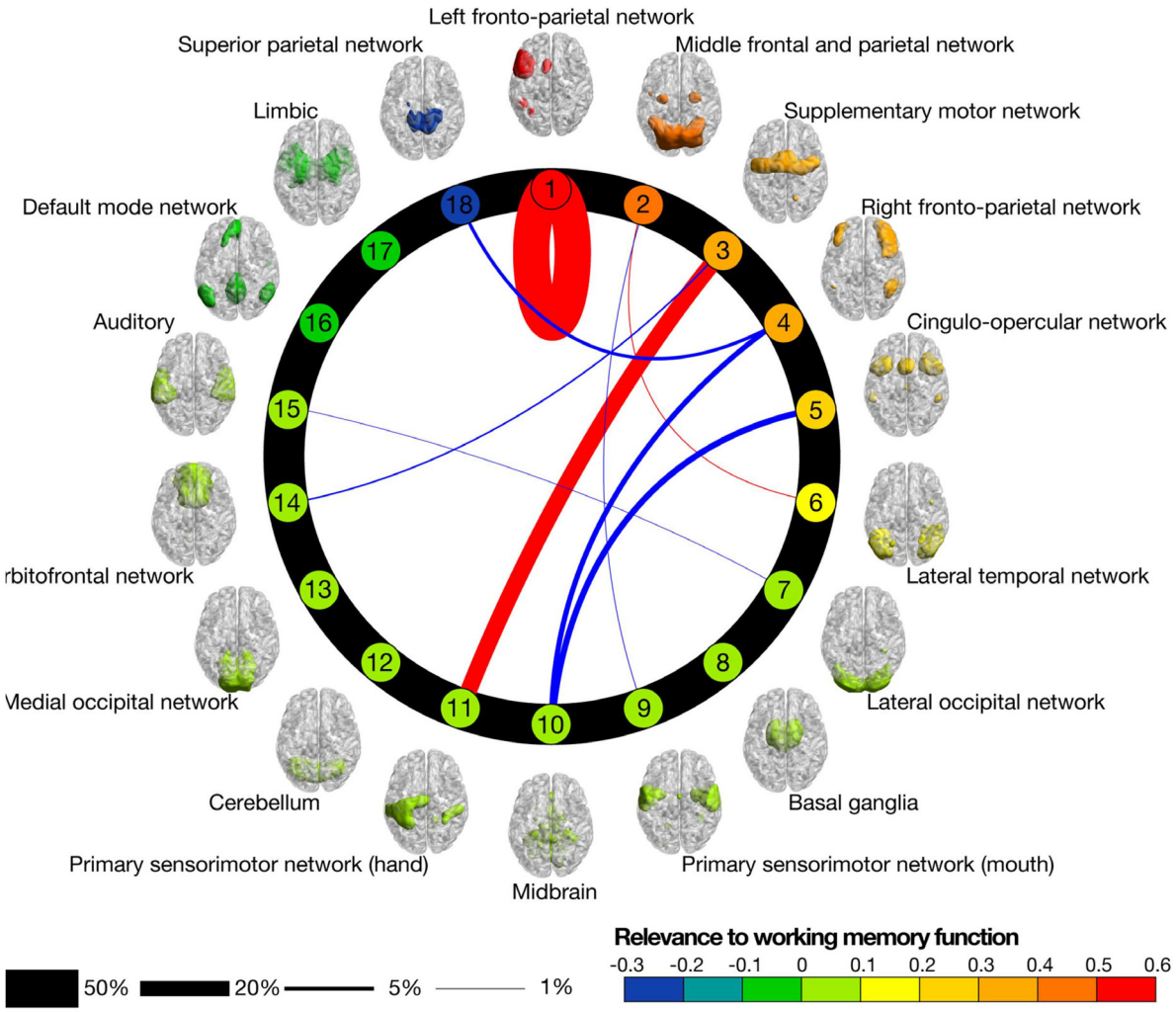
   Note that sleep research has long established the psychophysiological parellels between reported sleep states and EEG signatures. These show that the deep sleep associated with delta-wave activity (1–3 Hz) typically has little or no reportable conscious experience (Dement and Wolpert, 1958). The more rapid EEG activity normally associated with non-sleeping states qualifies as an NCC for C*. This result, however, illustrates the difference between an NCC and a NSCP, since the absence of delta waves does not qualify as a substrate despite correlating with positive C*, and the remaining EEG activity does not switch off during delta wave sleep.

   A prime example of empirical use of the extinguishability criterion is a study by Koubeissi et al. (2014), in which they found that electrical stimulation of the (left) **claustrum** above a certain threshold reversibly extinguished the participant's C* for the time period of the stimulation, whereas corresponding stimulation of nearby brain regions had no such effect. This result suggests that the claustrum is an important component of the NSCP, which must therefore have a spatially localized substrate at least including the claustrum. (They did not have access to

the right claustrum, so the respective roles of its two hemispheric partitions is undetermined).

5. **Iterativity.** Any plausible NSCP measure must exhibit the iterativity of repeated conscious experiences over the experiential range of time scales. This was the case for the electrical stimulation of the temporal lobe by Penfield (1958), where same long-forgotten conscious memory sequence was repeatedly evoked by stimulation at a single site in the temporal lobe (but not any other part of the brain). Without stimulation, the iterativity could be assessed by looking for long-range correlations within EEG or fMRI signals from memory areas, such as the temporal lobe. The intrinsic signal can be segmented into sliding segments each correlated with the next. Then the process is repeated at different scales of segment length. If any two segments show a significant correlation, this signal segment is then correlated throughout the signal duration to look for further repeats. In this way regular iterative patterns at a range of timescales can be uncovered.
6. **Operationality.** Operationality is the "working" aspect of working memory, the functionality of not only bringing relevant memories to C* but using them to answer questions either in relation to a single memory or in relation to the relationships between memories. This functional property is readily accessible to empirical techniques such as fMRI. Various NSCP brain sites associated with working memory have been identified through a vast range fMRI studies, but their specific roles and dynamics in relation to the operational properties of **experiential** working memory are not well understood. A particular example is the neural connectivity study of Yamashita et al. (2015), which assessed the intrinsic connectivity among 18 previously identified brain networks during learning of a three-back working-memory task: Is the current image the same as the one three-images back in the sequence? (which would be an example of a first-order question asked in relation to each individual memory). The performance improvement in this task was almost entirely attributable to the self-interaction of the dorsolateral prefrontal network (out of 171 possible network connections), with a few other weak contributions such as between primary and secondary motor cortex (see **Figure 2**). This result gives interesting insight into the improvement of operationality (and hence into the operationality *per se*) of working memory, with the effects localized to a particular cortical region that has been strongly associated with working memory in past studies.

   The operational property can be tested either behaviorally or physiologically by using a behavioral task that requires accessing an operational relationship between two previously unrelated memories. A prime example of such operations is the mathematical task of performing an arithmetic operation on two numbers of a form that is not pre-encoded by mathematical tables (i.e., not memorized). The participant thus has to perform the operation of a real-time calculation to solve the task, requiring the retention of the numbers and manipulating the intermediate solutions in working memory to complete the calculation (Metcalfe et al., 2013).
7. **Multifacetedness.** The test for the property of multifacetedness is that the neural activity proposed as the NSCP for C* should be activated for all the multifaceted aspects of experiential consciousness. The testability criterion would be that any measurable neural process identified as the NSCP would be concurrent with one such experience, and *vice versa*, with no significant misses or false alarms in the coupling instances. One form of such multifacetedness is provided by the network interaction study of Yamashita et al. (2015; **Figure 2**). The 18 specified networks each have a defined function in the mental C* lexicon, although only a few are identified by the authors. In any particular task, several or many of these networks may be expected to be activated, with this activation representing the degree of multifacetedness of the C* experience. The relation of the identified functions of the activated networks to the subjective reports of the performers of such tasks can provide an index of the degree to which the fMRI activations represent the NSCP. To avoid reporting bias, it may be necessary to provide the reporters with a list of possible functions corresponding to the brain networks, for them to assess the degree of expression of each of them (and others not on the list).
8. **Complex interconnectivity.** Any one aspect of consciousness, such as awareness of a face, is not a simple state but a multilevel complex of experiential components from the basic "raw feel" to the communicative socio-emotional implications. To be explanatory, the NSCP should exhibit a similar variety of interconnectivity. A probe for such interconnectivity is provided by the network interaction study of Yamashita et al. (2015; see **Figure 2**). Although they are termed "networks," most are dominated by one or a few cortical areas (in a 21st century manifestation of localization of function). Of the 171 possible connections, Yamashita et al. (2015) report strong connectivity among only nine of them (5%), predominantly motor networks, and negative associations between these networks and the Default Mode Network (DMN), whose function seems to be non-task-related personal reverie and planning. Thus, at this level of network analysis there is complex interconnectivity *within* each individual network, but relative isolation (specialization of function) *among* the networks.
9. **Autosuppressivity.** Once again, the attentional autosuppressivity that keeps C* moving on from each identifiable state to the next is a further property that can be identified in candidate mechanisms for the NSCP. This property is already well-substantiated as "inhibition of return" in attentional and saccadic target selection in visual search (Posner and Cohen, 1984; Rafal et al., 1989). The issue with such studies is that they do not index C* directly, so the inhibition of return is not necessarily associated with C* *per se*, in both directions of decoupling: inhibition of return is observed at levels of saccadic control that are not normally associated with C* (Posner et al., 1995); conversely, it is *not* observed in parallel search, where perceptual "popout" is a fully conscious



Network label	Brain regions
Left fronto-parietal network	left-lateralized fronto-parietal regions
Middle frontal and parietal network	middle frontal gyri and superior parietal lobules
Supplementary motor network	superior and middle frontal gyri including supplementary and premotor areas as well as frontal eye field
Right fronto-parietal network	right-lateralized fronto-parietal regions
Cingulo-opercular network	bilateral anterior insula/frontal opercula and the anterior aspect of the body of the cingulate gyrus
Lateral temporal network	middle and inferior temporal gyri
Lateral occipital network	lateral posterior occipital cortices
Basal ganglia	bilateral basal ganglia and thalamus
Primary sensorimotor network (mouth)	dorsal precentral gyri, central sulci, postcentral gyri, superior and inferior cerebellum
Midbrain	midbrain
Primary sensorimotor network (hand)	ventral precentral gyri, central sulci, postcentral gyri, superior and inferior cerebellum
Cerebellum	cerebellum
Medial occipital network	medial posterior occipital cortices
Orbitofrontal netwrok	subgenual anterior cingulate cortex and orbitofrontal cortex
Auditory	transverse temporal gyri
Default mode network	medial prefrontal and posterior cingulate/precuneus areas
Limbic	limbic and medial-temporal areas
Superior parietal network	superior parietal lobule

FIGURE 2 | The intrinsic connectivity study of Yamashita et al. (2015), showing the strength of functional connectivities among 18 brain networks during a working memory task.

phenomenon (Treisman and Gelade, 1980). However, the control mechanisms underlying such autosuppressivity do not have to be conscious for them to form the basis of the experienced autosuppressivity, and in order to form its NSCP they only need to operate when the autosuppressivity occurs. As such, the neural bases for inhibition of return remain strong candidates for the NSCP of this aspect of C*.

10. **Self-referentiality.** Computationally, it is not difficult to construct a computer program that includes itself as a component in its representation. Indeed, the representation of the external player as an element in the programmed domain is a common feature of computer games known as an "avatar." Such an avatar escapes Russell's Paradox by not being a full representation that actually contains itself, but only a reduced representation of the major features of itself in model form. It is not so clear how the neural implementation of an avatar could be achieved, but to do so is a further prerequisite of the NSCP. Note that this concept, of self-referentiality being a testable aspect of the NSCP while referentiality *per se* is not, is itself paradoxical. Self-referentiality can be tested by identifying a brain process that switches on and off concurrently with the switch between awareness of the self "avatar" and of other content, whereas referentiality cannot be tested because it is an unavoidable property of C*, and there is no non-referential form of C* against which to test the "off" state of a candidate process.

DISCUSSION

Functions of Consciousness

A further aspect of consciousness that can be considered is its **evolutionary function** (Bridgeman, 2011; Earl, 2014), as distinct from its neural integrative, adaptive and working memory functions, which are commonly highlighted (e.g., Baars et al., 2013). Indeed, many aspects of brain function are integrative, adaptive and mnemonic without passing the threshold of conscious awareness, such as the procedural memory functions of the cerebellum and basal ganglia. It is evident, therefore, that such brain functions do not require consciousness *per se*, and that such neural integrative, adaptive and mnemonic functions therefore do not require consciousness to operate.

Bridgeman (2011) argues that consciousness allows organisms to avoid the tyranny of response to the immediate (e.g., Pavlovian) environment, allowing the organism to superpose its goal-directed needs into the situational response. He concludes that such behavior requires the operation of working memory, and that consciousness is therefore a particular form of working memory. However, although **goal-directedness** may be a characteristic property of consciousness, it does not seem a sufficient criterion for the inference of consciousness as an **experience**. Virtually all behaving organisms engage in such goal-directed behavior in one form or another, but we would hesitate to ascribe consciousness to all forms of goal-directed behavior (such as cows eating grass, for example). Indeed, goal-directed behavior can be observed in single-celled micro-organisms, such as the hunting behavior of dinoflagellates and planaria, based on the information gleaned from their unitary subcellular eye (Schwab, 2012). Thus, behavioral goal-directness is a property – indeed, the essential property – of all behaving organisms, or animals, making it difficult to distinguish the role of goal-directedness in consciousness *per se* from that in behavior in general.

An alternative view of the role of consciousness in working memory is that it represents the **interface** of the memory storage process. There is substantial evidence that we can only remember items from the sensory world that were attended (i.e., that were a focus of conscious awareness; Penfield, 1958). Although unattended items may be processed in some form to allow their characterization as uninteresting targets for attention, through what is known as pre-attentive processing (Neisser, 1967), such items do not reach the site of accessible memory. Only attended items can be recalled from memory. It therefore seems that consciousness may represent the gateway to memory. While not all items that reach consciousness may be remembered, it seems to be the case that all items that are remembered must have reached consciousness.

Although consciousness is thus a *sine qua non* for laying down the memory for an item, it is nevertheless not required for the memory *per se*. Indeed, the very concept of memory implies a lack of consciousness, for the act of **remembering** corresponds precisely to bringing the item back into consciousness from its latent storage status outside of consciousness. This lack consciousness is evident for the vast range of items in long-term memory, such as the name of your first-grade teacher (which you may not have brought to consciousness for decades), but is also true for short-term iconic memory. We have all had the experience of being told a phone number, then doing competing activity during which we are not conscious of the number, then being able to recall the phone number by directing attention to the internal auditory "echo" of that number that is still available for a few minutes, though outside the immediate consciousness until it is accessed.

Equally, consciousness may be distinguished from the more interactive concept of **working memory** (Baddeley and Hitch, 1974), the earlier form of the global workspace that is currently being championed by Baars et al. (2013). There are three aspects to these forms of operation, which form the core operations of the process we call "thinking": the recall of items from memory, the sequence of working operations on the items, and the consciousness of this process. Consider the quiz question of whether an item is bigger than a breadbox, for an item such as a rugby ball. We have to recall the item from memory, examine the memory to ascertain its dimensions, do likewise for the standard concept of a breadbox, compare the sets of dimensions (with appropriate rotation to the best fitting orientation) and make the decision as to which is larger. Indeed, we have to decide which form of breadbox is intended, the single-loaf kind that would be too small for the rugby ball or the multiloaf breadbin that would easily

be large enough. Since each of these operations require the act of recall from memory, followed by operations on them, it seems to be a misnomer to call them a form of memory *per se*, even if it were an active form. The term "working memory" was perhaps a strategy to avoid the use of the term "consciousness" in the reductionist milieu of the mid-20th century, but including all these operations seems to inflate the storage function of memory to an implausible extent. It is preferable to restrict the term "memory" to the storage function of retaining the information after moving it outside the theater of consciousness.

Finally, what light does this analysis shed on the Baars' **global workspace** as the essence of consciousness? In the breadbox quiz, we become conscious of posing the question of recalling the memory of the rugby ball, of its scale, of recalling the breadbox, of its relative scale, of aligning the two up for comparison, and of the decision. But we are not conscious of all these factors at the same time. At least at the beginning of the process, while we are recalling the shape of the rugby ball from those of other sports balls, we are not conscious of considering the type of breadbox. It is only when all the components have been recalled from memory that we may perhaps be conscious of them all together. So, while the global workspace may be the specific arena of the **operations** of consciousness, it does not seem to be an accurate characterization of the core function of consciousness *per se*. Consciousness seems to be better characterized as the role of operational attention *within* the global workspace, rather than the global workspace as a whole. In this sense, consciousness is conceptualizable as a "mental grasping" capability requisite for the manipulation of mental constructs within the global workspace.

In summary, the evolutionary function of consciousness may be not so much a mechanism to introduce goal-directed aspects into the control of behavior as one to function as the gatekeeper for memory storage, such that only aspects of the sensory input that pass the criterion for reaching consciousness can be stored in memory, while all other aspects are lost (Penfield, 1958). The stored memories themselves decay over time, so they may also tend to be lost eventually, but many are retained for long periods, or even a lifetime, especially those that were experienced with heightened consciousness. Thus, while "attention" describes the selective function of which aspects of the sensory input are the focus of the gatekeeping function, "consciousness" describes the activation level through which the elaborated sensory input becomes laid down as a memory trace, and reinforced or reorganized in memory when recalled through the working memory mechanism.

Comment on Integrated Information Theory

Perhaps the most salient current analysis of consciousness is the Integrated Information Theory of Tononi (2008). The present analysis does not extend to a full evaluation of its claims, but it is relevant to address one its core axioms. This is the axiom that consciousness has a "rich conceptual structure composed of a very large number of concepts and relations" (Tononi et al., 2016, p. 457), which correspond to all the **phenomenal** distinctions that make up our reported conscious experience. To treat this property as axiomatic of consciousness seems to completely miss the point, however, since even a consciousness that is limited to very few concepts should still qualify as a valid form of phenomenal consciousness. Tononi's specification is roughly equivalent to the multifacetedness property of consciousness in the present analysis, so to that extent we agree, but to treat it as an axiomatic defining property of consciousness seems misguided (compare Bayne, 2018). Even if the maximum capacity of consciousness was severely limited, as it presumably must be in the lowest level of organism that experiences it, that limitation does not detract from the fact of that consciousness. Indeed, it is a common experience that one's consciousness becomes drasticaly limited in "conceptual structure" when one is very tired or otherwise debilitated, though it may still have the qualitative vividness that is the core characteristic of phenomenal consciousnees. Conversely, even if a complex system, a deep-learning computer or the Internet, develops a "very large number of concepts and relations," that does not mean that it is conscious. In this sense IIT (Tononi, 2008; Koch et al., 2016; Tononi et al., 2016) cannot be considered to be a theory of phenomenal consciousness *per se*, though it could be considered to be a valid conceptualization for what is termed "access consciousness."

CONCLUSION

This article has had the goal of expanding the soup-to-nuts philosophy of Emergent Aspect Dualism to the experienced properties of consciousness, as one of the prime forms of emergence, and one to which the only access is subjective report. To extend the probing to the NSCP, a full specification of the properties of consciousness as subjectively experienced is provided in forms that are neuroscientifically testable. These properties are then considered against those of the global workspace and IIT conceptualizations of consciousness to highlight the differences between those viewpoints and the current framework, which is the explicit testability of consciousness conceived as the experiential focus of operational attention by which transient sensory input is converted to long-lasting memories.

AUTHOR CONTRIBUTIONS

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

ACKNOWLEDGMENTS

I thank the two supportive reviewers for their analytic critiques and insightful suggestions.

REFERENCES

Baars, B. J. (1983). "Conscious contents provide the nervous system with coherent, global information," in *Consciousness and Self-Regulation* (Boston, MA: Springer), 41–79.

Baars, B. J., Franklin, S., and Ramsoy, T. Z. (2013). Global workspace dynamics: cortical "binding and propagation" enables conscious contents. *Front. Psychol.* 28:200. doi: 10.3389/fpsyg.2013.00200

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

Bayne, T. (2018). On the axiomatic foundations of the integrated information theory of consciousness. *Neurosci. Conscious.* 4:niy007. doi: 10.1093/nc/niy007

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

Bohm, D. (2002). *Wholeness and the Implicate Order (Online-Ausg. ed.)*. Hoboken: Routledge.

Bridgeman, B. (2011). Functions of consciousness. *Cogn. Neurosci.* 2, 115–116.

Chalmers, D. J. (1997). *The Conscious Mind in Search of a Fundamental Theory (Philosophy of Mind Series)*. Oxford: Oxford University Press, USA.

Charpak, C., Broch, H., and Holland, B. K. (2004). *Debunked! ESP, Telekinesis, and Other Pseudoscience*. Baltimore, MD: Johns Hopkins University.

Crick, F., and Koch, C. (1990). Towards a neurobiological theory of consciousness. *Semin. Neurosci.* 2, 263–275.

Davey, B., and McBride, S. (1986). Effects of question-generation training on reading comprehension. *J. Educ. Psychol.* 78:256. doi: 10.7334/psicothema 2013.252

Dement, W., and Wolpert, E. A. (1958). The relation of eye movements, body motility, and external stimuli to dream content. *J. Exp. Psychol.* 55, 543–553. doi: 10.1037/h0040031

Duan, L., Dai, R., Xiao, X., Sun, P., Li, Z., and Zhu, C. (2015). Cluster imaging of multi-brain networks (CIMBN): a general framework for hyperscanning and modeling a group of interacting brains. *Front. Neurosci.* 9:267. doi: 10.3389/fnins.2015.00267

Earl, B. (2014). The biological function of consciousness. *Front. Psychol.* 5:697. doi: 10.3389/fpsyg.2014.00697

Feynman, R. (1985). *QED: The Strange Theory of Light and Matter*. Princeton University Press.

Finster, F., Smoller, J., and Yau, S.-T. (1999). Particle-like solutions of the Einstein-Dirac equations. *Phys. Rev D* 59:104020.

Gazzaniga, M. (1985). The social brain. *Discov. Netw. Mind* 77–79.

Hartmann, R., Singh, P. K., Pearce, P., Mok, R., Song, B., Díaz-Pascual, F., et al. (2019). Emergence of three-dimensional order and structure in growing biofilms. *Nat. Phys.* 15, 251–256. doi: 10.1038/s41567-018-0356-9

Hasker, W. (1999). *The Emergent Self*. New York: Cornell University Press.

Huth, A. G., Lee, T., Nishimoto, S., Bilenko, N. Y., Vu, A. T., and Gallant, J. L. (2017). Decoding the semantic content of natural movies from human brain activity. *Front. Syst. Neurosci.* 7:81. doi: 10.3389/fnsys.2016.00081

Koch, C. (2004). *The Quest for Consciousness: A Neurobiological Approach*. Englewood: Roberts & Company Publishers.

Koch, C., Massimini, M., Boly, M., and Tononi, G. (2016). Posterior and anterior cortex - where is the difference that makes the difference? *Nat. Rev. Neurosci.* 17:666. doi: 10.1038/nrn.2016.105

Koubeissi, M. Z., Bartolomei, F., Beltagy, A., and Picard, F. (2014). Electrical stimulation of a small brain area reversibly disrupts consciousness. *Epilepsy Behav.* 37, 32–35. doi: 10.1016/j.yebeh.2014.05.027

MacKay, D. M. (1960). On the logical indeterminacy of a free choice. *Mind* 69, 30–40.

Metcalfe, A. W., Ashkenazi, S., Rosenberg-Lee, M., and Menon, V. (2013). Fractionating the neural correlates of individual working memory components underlying arithmetic problem-solving skills in children. *Dev. Cogn. Neurosci.* 6, 162–175. doi: 10.1016/j.dcn.2013.10.001

Minev, Z. K., Mundhada, S. O., Shankar, S., Reinhold, P., Gutiérrez-Jáuregui, R., Schoelkopf, R. J., et al. (2019). To catch and reverse a quantum jump mid-flight. *Nature* 570, 200–204. doi: 10.1038/s41586-019-1287-z

Müller, N. G., and Kleinschmidt, A. (2007). Temporal dynamics of the attentional spotlight: neuronal correlates of attentional capture and inhibition of return in early visual cortex. *J. Cogn. Neurosci.* 19, 587–593. doi: 10.1162/jocn.2007.19.4.587

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

Neisser, U. (1967). *Cognitive Psychology*. New York, NY: Appleton-Century-Crofts.

Penfield, W. (1958). Some mechanisms of consciousness discovered during electrical stimulation of the brain. *Proc. Natl. Acad. Sci. U.S.A.* 44, 51–57. doi: 10.1073/pnas.44.2.51

Penrose, R. (1995). *Shadows of the Mind: A Search for the Missing Science of Consciousness*. Oxford: Oxford University Press.

Posner, M., Rafal, R. D., Choate, L., and Vaughn, J. (1995). Inhibition of return: neural basis and function. *Cogn. Neuropsychol.* 2, 211–228.

Posner, M. I., and Cohen, Y. (1984). "Components of visual orienting," in *Attention and Performance X: Control of Language Processes*, eds H. Bouma and D. Bouwhuis (Hillsdale, NJ: Erlbaum), 531–556.

Rafal, R., Calabresi, P., Brennan, C., and Sciolto, T. (1989). Saccade preparation inhibits reorienting to recently attended locations. *J. Exp. Psychol.* 15, 673–685. doi: 10.1037//0096-1523.15.4.673

Rosenshine, B., Meister, C., and Chapman, S. (1996). Teaching students to generate questions: A review of the intervention studies. *Rev. Educ. Res.* 66, 181–221.

Schrödinger, E. (1935). Die gegenwärtige Situation in der Quantenmechanik (The present situation in quantum mechanics). *Naturwissenschaften* 23, 807–812.

Schwab, I. (2012). *Evolution's Witness: How Eyes Evolved*. New York, NY: Oxford University Press.

Searle, J. (1990). Who is computing with the brain? *Behav. Brain Sci.* 13, 632–642.

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

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

Treisman, A., and Gelade, G. (1980). A feature-integration theory of attention. *Cognit. Psychol.* 12, 97–136.

Turkheimer, F. E., Hellyer, P., Kehagia, A. A., Expert, P., Lord, L. D., Vohryzek, J., et al. (2019). Conflicting emergences. Weak vs. strong emergence for the modelling of brain function. *Neurosci. Biobehav. Rev.* 99, 3–10. doi: 10.1016/j.neubiorev.2019.01.023

Tyler, C. W. (2015). The emergent dualism view of quantum physics and consciousness. *Cosmos Hist.* 11, 97–114.

Tyler, C. W. (2018). The emergent aspect dualism view of quantum physics: a new ontology to resolve the complementarity conundrum. *J. Res. Philos. Hist.* 1, 166–182. doi: 10.22158/jrph.v1n2p166

Tyler, C. W. (2019). Points of contact between the Stappian philosophy and Emergent Aspect Dualism. *Activ. Nervosa Superior* 61, 6–11.

Vegt, W. (2019). Unification of the Maxwell-Einstein-Dirac correspondence. *FrenXiv* [Preprint]. doi: 10.31226/osf.io/qrpft

Wigner, E. P. (1970). Physics and the explanation of life. *Found. Phys.* 1, 35–45. doi: 10.1007/bf00708653

Wittgenstein, L. (1953). *Philosophical Investigations*. Blackwells: Oxford.

Yamashita, M., Kawato, M., and Imamizu, H. (2015). Predicting learning plateau of working memory from whole-brain intrinsic network connectivity patterns. *Sci. Rep.* 5:7622. doi: 10.1038/srep07622

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

HYPOTHESIS AND THEORY
published: 25 June 2020
doi: 10.3389/fpsyg.2020.01399

A Phenomenological Paradigm for Empirical Research in Psychiatry and Psychology: Open Questions

Leonor Irarrázaval[1,2]*

[1]Section Phenomenological Psychopathology and Psychotherapy, Psychiatric Department, University Clinic Heidelberg, Heidelberg, Germany, [2]Centro de Atención Psicológica, Facultad de Ciencias Sociales y Humanidades Sede Talca, Universidad Autónoma de Chile, Talca, Chile

This article seeks to clarify the way in which phenomenology is conceptualized and applied in empirical research in psychiatry and psychology, emphasizing the suitability of qualitative research. It will address the "What," "Why," and "How" of phenomenological interviews, providing not only preliminary answers but also a critical analysis and pointing to future directions for research. The questions it asks are: First, what makes an interview phenomenological? What are phenomenological interviews used for in empirical research in psychiatry and psychology? Second, why do we carry out phenomenological interviews with patients? Is merely contrasting phenomenological hypotheses or concepts enough to do justice to the patients' involvement? Third, how should we conduct phenomenological interviews with patients? How can we properly perform analysis in empirical phenomenological research in psychiatry and psychology? In its conclusion, the article attempts to go a step beyond these methodological questions, highlighting the "bigger picture": namely, the phenomenological scientific paradigm and its core philosophical claim of reality as mind-dependent.

Keywords: applied phenomenology, methodology, qualitative research, psychiatry and psychology, phenomenological interviews

OPEN ACCESS

Edited by:
Anthony Vincent Fernandez, Kent State University, United States

Reviewed by:
Zeno Van Duppen, University Psychiatric Center KU Leuven, Belgium
Kristian Moltke Martiny, University of Copenhagen, Denmark

***Correspondence:**
Leonor Irarrázaval
leonor.irarrazaval@uautonoma.cl

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

Received: *18 September 2019*
Accepted: *25 May 2020*
Published: *25 June 2020*

Citation:
Irarrázaval L (2020) A Phenomenological Paradigm for Empirical Research in Psychiatry and Psychology: Open Questions. Front. Psychol. 11:1399. doi: 10.3389/fpsyg.2020.01399

INTRODUCTION

An initial proposal in favor of "naturalizing phenomenology" was presented in the article "First-person methodologies: What, Why, How?" published by Varela and Shear (1999) in the *Journal of Consciousness Studies*. The authors were not only concerned with the need for a method in cognitive sciences to obtain empirically-based descriptions of the subject, but also with providing the basis for a "science of consciousness." "Neurophenomenology" was proposed by Varela (1996) as a means of linking first- and third-person perspectives through a systematic examination of subjective experience within experimental settings. An important requirement of neurophenomenology was that both experimenter and experimental subject must learn the Husserlian phenomenological method. The notion "phenomenology" was employed in the etymological sense of the term, that is, "the study of that which appears" (from Greek *phainómenon* "that which appears" and *lógos* "study"). Additionally, Varela (1990) coined the term "enactive," meaning not to act out or to perform as on a stage, but to "enact," that is, "to bring forth" or to "emerge" (*hervorbringen*, in German), as it is used in the phenomenological tradition.

Accordingly, the phenomenological method was conceived and applied as a form of training one's attention to that which "appears" in the subject's conscious experience, making it similar to a meditation technique. Examples of neurophenomenology are the experiments led by Lutz et al. (2002), which analyzed subjective reports, reaction times, and brain activity. However, a different approach was proposed by Gallagher (2003), who claimed that a "phenomenologically enlightened experimental science" means incorporating concepts and distinctions from the phenomenological analysis into the actual design of an experiment. In contrast to neurophenomenology, this approach does not require learning the Husserlian phenomenological method or even making first-person reports in the experiments. Examples of "front-loaded phenomenology" are neuroimaging experiments employing the phenomenological distinction between "sense of agency" and "sense of ownership" in involuntary movement (Ruby and Decety, 2001; Chaminade and Decety, 2002; Farrer and Frith, 2002).

However, experimental designs are normally not classified as part of qualitative research methodologies (Fischer, 2006; Maxwell, 2011, 2012; Patton, 2015; Creswell and Poth, 2018). One of the clearest differences between qualitative and quantitative approaches is that qualitative research is carried out in everyday natural conditions, rather than in experimental settings. Concerning the qualitative/quantitative distinction, there is an ongoing debate not only around the differences between the two approaches (Morgan, 2018; Maxwell, 2019), but also around whether they are actually distinguishable at all (Hammersley, 2018). Whatever their differences or similarities, qualitative and quantitative approaches are commonly conceived as compatible and their integration – in the form of mixed-methods research designs – valuable (Tashakkori and Teddlie, 2010). So, the incorporation of phenomenological interviews in experimental designs is one kind of mixed-method research design: One example is neurophenomenology, where the qualitative component is provided by phenomenology. Broadly speaking, qualitative research is used in many social sciences and humanities disciplines, including psychology, sociology, political sciences, and anthropology. A range of techniques are employed in qualitative research to gather experiential data, such as open-ended interviews, direct observation, focus groups, and document analysis (e.g., clinical records and personal diaries), and different methods are used for the associated qualitative data analysis, including phenomenology, ethnography, narrative analysis (e.g., biographical and life story studies), case studies, and grounded theory. In contrast to the large sample sizes needed in quantitative research to accomplish statistical validation of the results, qualitative research is characterized by an in-depth approach, which means working with few cases, with representativeness not being of such key importance (Barbour and Barbour, 2003). The use of less structured methods allows for the emergence of ideographic descriptions, personal beliefs and meanings, thus addressing the experiential processes of the subjects being studied (Schwartz and Jacobs, 1979; Barbour, 2000; Maxwell, 2011, 2012, 2019).

This article shall not focus on experimental phenomenology. However, this is no way meant to discredit in any sense this form of research design. Indeed, mention has already been made of the precursors of the experimental application of phenomenology to acknowledge the important contribution this research tradition has made – and continues to make – in ensuring that phenomenology acquires a scientific status. For instance, the project "cardiophenomenology" has been recently proposed by Depraz and Desmidt (2019) as a refinement of Varela's neurophenomenology and performed in experimental studies of surprise in depression (Depraz et al., 2017). In addition, it is worth mentioning Martiny's (2017) transdisciplinary research on the phenomenological and neurological aspects of living with brain damage, specifically cerebral palsy. Martiny's work not only has been influenced by, but also seeks to revitalize, Varela's "radical" proposal, reminding us of the importance of working with openness and a change of mindset in cognitive science. Usually framed as "embodied cognition," this proposal approaches the mind as embodied, embedded, enacted, and extended (4E cognition), implying an awareness regarding the fact that the "embodied" notion applies not only to the mind of the experimental subject but also to the cognitive scientist carrying out the research (Depraz et al., 2003). Indeed, phenomenology has breached the frontiers of the philosophical discipline to influence the development of interdisciplinary fields of studies bridging the biomedical sciences and the humanities. Besides its application in the cognitive sciences, phenomenology is currently being widely applied in empirical research in healthcare-related disciplines, mostly in psychiatry and psychology. The most influential empirical application of phenomenology has been in the field of psychopathology, with the development of phenomenological interviews for the investigation of schizophrenia spectrum disorders (Parnas et al., 2005; Sass et al., 2017). However, the extent of phenomenology's applicability outside the strict domain of philosophy is currently a topic of intense debate and controversy (Zahavi and Martiny, 2019). The conceptualization of phenomenology in the literature of qualitative research, which has been mostly developed in North America, is not always in line with that of the continental European philosophical tradition. Recent years have seen the start of a dialogue bridging the two traditions, qualitative research and philosophical phenomenology, giving a promise of fruitful collaboration in the future.

This article will address the "What," "Why," and "How" of phenomenological interviews, reviewing recent empirical research in the field of phenomenological psychopathology and psychotherapy. Important to note is that qualitative research, as described above, refers to empirical research, not to basic or theoretical investigations. Phenomenological qualitative research in psychology has been developed using Husserlian concepts such as the *"epoché"* and the "phenomenological reduction," and precisely on the use of such conceptualizations is where most of the current discussion has been placed. The article, therefore, will not attempt to provide a broad understanding of the phenomenological tradition. Instead, it will focus on a more specific discussion of methodological issues concerning the empirical application of phenomenology in qualitative research in psychiatry and psychology, and Husserl's methodology in particular. To do so, we first need to agree that the application

of phenomenology to empirical research in psychiatry and psychology employing interviews is qualitative, not quantitative. In a strict sense, quantitative methodology based on frequency and scales of severity of the patients' anomalous experience, although necessary for the statistical validation of the interviews, goes beyond the scope of phenomenology. According to the phenomenological approach, mental disorders cannot be reducible to a cerebral organic basis, nor to numbers, as they are not entities *per se* but psychopathological configurations that can be identified in the diagnostic process of interaction between a clinician and a patient (Fuchs, 2010a; Pallagrosi et al., 2014; Pallagrosi and Fonzi, 2018; Gozé et al., 2019). Consequently, phenomenological interviews are designed to address not objective, but subjective data, namely the *what it is like* of patients' anomalous experiences. In this way, the patients' descriptions of their subjective experiences are not conceived as "static" entities, but, rather, as part of dynamically, open-ended developing processes and interpretations (Martiny, 2017).

WHAT

What makes an interview "phenomenological"? What are phenomenological interviews used for in empirical research in psychiatry and psychology?

Medical psychiatric diagnosis relies on standardized manuals providing a description of the apparent symptomatology and mostly excludes any assessment of subjective experience (Mishara, 1994; Parnas and Zahavi, 2002; Fuchs, 2010a). Under this approach, research in psychiatry has mainly developed from a third-person perspective, using the methods of the physical and natural sciences. Biomedical psychiatry has prioritized the use of quantitative methods and statistical analysis, whereas the value of qualitative in-depth analysis has been underestimated. The preferred experimental design has been the randomized controlled trial to demonstrate the efficacy of treatments involving psychoactive drugs (Deacon, 2013; Deacon and McKay, 2015). An alternative conceptual model to this comes from the phenomenological tradition of psychopathology. In order to understand and conceptualize the anomalous experience of a given mental illness, the phenomenological diagnosis highlights the importance of assessing patients' subjectivity. Over the last two decades, phenomenological interviews have been developed to complement standardized diagnostic systems such as Diagnostic and Statistical Manual of Mental Disorders (DSM-5) (American Psychiatric Association, 2013) and International Statistical Classification of Diseases and Related Health Problems (ICD-10) (World Health Organization, 2012). The most important phenomenological interviews are the Examination of Anomalous Self-experience (EASE, Parnas et al., 2005) and its supplement, the Examination of Anomalous World Experience (EAWE, Sass et al., 2017). These interviews have been inspired by the Husserlian tradition and have incorporated classical descriptions of phenomenological psychopathology (particularly from Blankenburg, Conrad, and Minkowski, among other authors). Their semi-structured design allows for an in-depth examination of the patients' subjective experiences within formal structures, such as corporeality, temporality, spatiality, and intersubjectivity. In this way, the descriptive task is not carried out on a totally random basis, as the interviews have specific domains and items that have already been established to guide the examination of the patient's experience. EASE and EAWE were developed with the chief purpose of exploring and better understanding patients' experiential and behavioral manifestations of schizophrenia spectrum disorders. These interviews offer comprehensive descriptions of disorders of the pre-reflexive self or *ipseity* (Sass, 1992; Parnas and Handest, 2003; Sass and Parnas, 2003; Parnas and Sass, 2008; Raballo et al., 2009; Fuchs, 2010b, 2013a; Sass et al., 2018). Indeed, EASE and EAWE have had great international impact in clinical practice and empirical research in psychiatry and psychology, and EASE has been translated into more than 10 languages, among them German, Danish, Spanish, Italian, and French.

EASE and EAWE describe aspects of the patients' anomalous experience that are not only relevant for diagnostic but also for psychotherapeutic purposes, as they can be useful as tools in both psychotherapeutic settings and in psychotherapy research. However, phenomenological psychopathology has focused primarily on the issue of psychiatric diagnosis, while the treatment of mental illness has remained less developed. Only in recent years has the treatment of mental illness become the focus of stronger research interest, directly involving the practice of psychotherapy (Fuchs et al., 2019). For its part, although not rooted in phenomenology, body-oriented therapy has been linked to a phenomenological framework, as it provides empirical evidence for embodiment-approach conceptualizations (Fuchs 2005; Fuchs and Schlimme, 2009; Koch and Fuchs, 2011; Fuchs and Koch 2014). The embodiment approach regards schizophrenia as a fundamental disturbance of embodiment, namely a "disembodiment," that entails a diminishment of the basic sense of self, a disruption of implicit bodily functioning and, as a result, a disconnection from intercorporeality with others. A range of empirical research into body-oriented therapy has been carried out in the field of phenomenological psychopathology. Empirical evidence of the effectiveness of body-oriented therapy for schizophrenia has been obtained from quantitative research carried out with manualized interventions (Röhricht and Papadopoulos, 2010) and using randomized controlled trials to measure outcomes (Martin et al., 2016). Recent research has incorporated phenomenological interviews to describe therapeutic change processes in body-oriented therapy for schizophrenia, thus explaining the relationship between processes and outcomes (Galbusera et al., 2018). Unsurprisingly, the phenomenological interviews revealed an understanding of change as a recovery of a "sense of self" in patients with schizophrenia (Galbusera et al., 2019).

The conceptualization of schizophrenia as a disorder of the self is shared by a number of philosophical and clinical approaches: it is not exclusive to phenomenological psychiatry (Parnas and Henriksen, 2014). So, in much the same way as body therapy has been "converted" to phenomenology, any other psychotherapeutic approach might well incorporate "front-loaded phenomenology," in the sense of the possibility of being linked to the phenomenological framework. This is especially the case when the effectiveness of psychotherapy

has been widely evidenced and recognized independently of its theoretical framework (Campbell et al., 2013). For instance, narrative/dialogical psychotherapy addressing schizophrenia as a disorder of the self might be consistent with the phenomenological conceptualization and could even serve as a complement for body-oriented therapy. In fact, EASE's and EAWE's rich descriptions provide evidence that patients with schizophrenia are able to communicate their experience in a comprehensive narrative form, which is quite contrary to Martin et al.'s (2016) claim that verbal dialogue can be difficult in patients with severe mental disorders. A suitable alternative might be the "metacognitive model" (Lysaker et al., 2018a). Under this model, deficits in metacognition undermine the availability of a sense of self, others, and the world, making it difficult to provide an adequate response to everyday-life situations. To deal with this, the so-called metacognitive reflection and insight therapy (MERIT) has been designed to target metacognition and recover the availability of a sense of self in the patients' experience (Lysaker et al., 2018b). Precisely because contemporary phenomenological psychiatry places particular emphasis on the bodily and pre-reflective level of experience, the use of phenomenological interviews to explore change process in MERIT might reveal interesting relationships between pre-reflexive and reflective forms of self-experience.

Does psychotherapy needs be rooted in the phenomenological tradition in order to be called "phenomenological?" Here we are talking about enterprises such as Freud's psychoanalysis or Binswanger's existential analysis/daseinsanalysis. Such an enterprise requires a well-achieved and comprehensive conceptualization of phenomenological psychopathology as well as a consequent psychotherapeutic intervention rooted in the same phenomenological conceptualization. Certainly, psychotherapy does not need to be rooted in phenomenology, although this enterprise, not a minor one, might be worth undertaking. Yet, the very essence of phenomenological psychotherapy is to remain faithful to the patient's self-experience and their constitutive vulnerability (Fuchs, 2013b; Irarrázaval, 2013, 2018; Irarrázaval and Sharim, 2014; Škodlar and Henriksen, 2019). Consequently, the development of integrative models of psychotherapy both bodily and narrative/dialogical addressing the patients' experience of vulnerability is definitely a future challenge.

WHY

Why do we carry out phenomenological interviews with patients? Is merely contrasting phenomenological hypotheses or concepts enough to justify the patients' involvement?

The justification for empirical research employing phenomenological interviews is extremely important, especially when persons with mental illnesses are involved. It is not only a matter of gathering data from the patients' experience but also one of what to do with this data and, in the end, what for. It is an ethical issue concerning the impact phenomenological interviews might have on patients interviewed. Any interview aimed at exploring the experience of a patient always involves some kind of intervention, so even when applied by accredited experienced clinicians, an ethical justification is required. Arguments before ethics committees that phenomenological interviews are beneficial and do not worsen patients' instability need to be convincing. Recalling and enacting in patients disturbing experiences we aim to grasp is certainly an intervention that needs justification. Obviously, phenomenological interviews are not psychotherapeutic interventions in themselves – that is, the dialogue in psychotherapy is not an interview – but they can be justified on the grounds similar to those usually employed by psychotherapy: the possibility of sharing anomalous experiences through an accepting and understanding communication helps patients to recover a sense of familiarity with their experience, thus reducing their sense of self-alienation. Furthermore, by means of the descriptive tasks called for in the semi-structured interviews, patients improve their articulation of anomalous experiences, which might have been otherwise overlooked, neglected, or even remain ineffable for them (Zahavi and Martiny, 2019).

Phenomenological interviews have been simply defined as falling within the framework of an interview "which is informed by insights and concepts from the phenomenological tradition and (which) in turn informs a phenomenological investigation" (Høffding and Martiny, 2016, p. 540). However, phenomenological interviews involving patients with mental illness should not only be consistent with insights and concepts from the phenomenological tradition of philosophy and psychopathology but, most importantly, they must make explicit their contribution to both diagnosis and psychotherapy. While a biomedical psychiatric diagnosis is ultimately oriented toward finding a suitable pharmacological treatment, a phenomenological diagnosis is ultimately oriented toward providing a treatment based on the experiential dimension of a given mental illness. The interest of a psychotherapist goes beyond the psychiatric diagnostic emphasis by approaching the patient as a whole person, aiming to understand the anomalies of experience within his/her social, cultural, and historical context. This broader, psychological, approach enables an understanding not only of how patients make sense of their anomalous experiences but also of how symptoms manifest themselves within the patients' immediate life context, as well as how a certain mental illness configures itself along the patients' history of meaningful interactions with others (Irarrázaval and Sharim, 2014; Irarrázaval, 2018). However, in spite of the importance given to the analysis of the patients' biography by several authors from the phenomenological tradition of psychopathology (Jaspers, Binswanger, and Blankenburg, among other authors), "biographical methods," originally developed for sociological research in the influential "Chicago School" (Bornat, 2008), have not been sufficiently incorporated in current phenomenological empirical research in psychiatry and psychology.

HOW

How should we conduct phenomenological interviews with patients? How can we properly perform analysis in empirical phenomenological research in psychiatry and psychology?

A phenomenological interview involves a second-person situation, in which the dialogical communication with the patient is crucial. No matter how strange or unrealistic the patients' anomalous experiences might appear to the interviewer, an attitude of professional competence and familiarity is necessary (Nordgaard et al., 2013). For the patient, anomalous experiences are actually lived experiences despite their lack of commonsensical validity. Hallucinations and delusions are, like nonpsychotic experiences, first-personally given, which means that they have a solipsistic validity. This is one of the reasons why it is difficult, especially in psychotic phases, for patients to come to terms with the fact that what they actually experience is not credible or real in the eyes of others, and even abnormal or pathological in the eyes of the clinician. Clearly, the interviewer's role is not to confront or contradict this lack of commonsensical validity, but simply to grasp the experiences as they appear to the patients. In other words, the interviewer conducts the interview with an attitude of empathetic understanding. Empathy should not be reduced to an attempt to understand the patient in a "representational" manner, in the sense that it does not refer to the interviewer's own experience of processing (imitating, thinking, or imagining) the patient's subjectivity (Irarrázaval, 2019). Empathy is the condition of possibility for the "subject-subject" relationship (Zahavi, 2015). That is to say, empathy is a distinct mode of other-directed intentionality that permits the unfolding of the patient's experience, approached as a unique other person. In this sense, empathic understanding permits the unfolding of the *what it is like* of the patient's anomalous experience.

In phenomenological interviews, why-like questions lead patients to respond with causal explanations of the anomalies of their experience or diagnosed mental illness, such as judgments, beliefs, theoretical constructions, etc., For their part, how-like questions guide patients to describe the way in which they live their experience, that is, the way in which the anomalies actually appear to the patients in their experience. To put it another way, both types of questions lead patients to talk about experiential contents, but in different ways: causal attributions in the former, and appearances in the latter. Causal attributions are by no means irrelevant aspects of the patient's experience not worth addressing in the interview. The way in which patients' attribute causes to their anomalous experience or mental illness can also provide valuable information for both diagnosis and psychotherapy. Moreover, the relationship between causal attributions and appearances is certainly valuable, as it entails a circular, dynamic process in which both orders of experiencing constantly influence one another. However, the gathering of phenomenological data is generally not aimed at obtaining causal explanations or attributional reports, as in the case of cognitive psychology, but mainly at exploring aspects of experience that how-like questions are designed to unfold.

Turning to data analysis, it has been said that phenomenology is interested in describing the formal structure of the experience rather than its content (Gallagher and Zahavi, 2008), but what does this actually mean? It seems difficult to imagine an experience as a mere structure without any content. Moreover, it is not possible to establish a category of experience that has not been previously built upon any content analysis. In qualitative studies, categories are built upon the basis of prior content analysis; both hypotheses and categories are developed as the study progresses and emerge from the data itself (Morrow, 2005; Maxwell, 2012), so-called "iterative process" (Barbour and Barbour, 2003). EASE and EAWE were built collecting first-person descriptions by a significant number of patients (around 100 each), which allowed for their statistical validation. However, only a fairly general description has been provided of how EASE's domains and items were developed: singular contents of anomalous experience are conceptualized and interconnected within a comprehensive system of meaningful structural wholes or *Gestalts*, leading to the "core" underlying psychopathological configuration (Nordgaard et al., 2013). A recent qualitative study on the responses to the two scales highlights the specificities of the phenomena described by EASE and EAWE, indicating that disturbances of world experience are fundamentally less unitary, while the experience of the self presents a more coherent and unitary *Gestalt* (Englebert et al., 2019).

Beyond the statistical validation of the interviews, replication is needed in other clinical samples and cultures to support previous findings and provide added evidence when compared with multiple clinical groups and cross culturally. However, if the focus of the analysis is placed merely on formal structural aspects, then when applying EASE and EAWE to new patients, we will not find domains or categories different from those already defined. To put it differently, quantitative replication of EASE or EAWE in other samples would barely lead to any new knowledge, because already established domains and items tend to constrain the patients' responses. So, particularly in terms of their potential contribution to psychotherapy, the best contribution that could be made from applying EASE and EAWE to new patients would result from a content analysis of the patients' reports. However, one key question concerning these interviews' replication remains unanswered: Which is the most appropriate qualitative method for analyzing the patients' descriptions?

The empirical application of Husserl's phenomenological method outside the strict scope of philosophy still is a topic of ongoing debate in both philosophy and the cognitive sciences. According to Zahavi (2019a,b,c), in philosophy, the main goal of phenomenology is not purely descriptive or attentive to how things appear to the subject; it focuses neither on the subject nor on the object, but on the correlation between them. In this context, the term *epoché* is used to refer to suspending or putting between parentheses a "naïve" or "natural" attitude toward reality in order to reflect upon fundamental ontological questions, thus adopting a critical stance on the conception of reality as mind-independently given. *Epoché*, usually described as putting "in brackets" the prejudices and theoretical assumptions of the interviewer (Fischer, 2009), in order to access phenomena as they appear in the subject's experience, has little to do with the original philosophical method. This does not imply that bracketing our prejudices and theoretical assumptions would not be desirable to avoid bias when conducting phenomenological interviews or analyzing

data (we can find several techniques for doing so). It is also not so important to calling such bracketing *epoché*, as long as we have a basic notion of Husserl's original sense of the term.

Phenomenology has been applied in empirical research not only in psychiatry and psychology, but also in other healthcare-related disciplines, such as nursing studies (Zahavi and Martiny, 2019). Nevertheless, the different forms in which phenomenology has been applied in these disciplines have been also controversial due to their divergence from the original Husserlian philosophical method (Zahavi, 2019b,d). For instance, some have questioned whether the method of analysis proposed by Giorgi (2009, 2012), "descriptive phenomenological psychological method," should be considered "phenomenological" or given another label. This method is aimed at the establishment of inclusive categories resulting from the content analysis of subjects' descriptions. In fact, Giorgi's method of content analysis seems closer to an adapted form of "eidetic variation" and quite different to the original Husserlian sense of the *epoché*, because it basically consists of summarizing the content of the interview transcript by deleting its redundancies, in order to reveal invariables or essences in "meaning" (see Irarrázaval, 2015). Eidetic variation is a conceptual analysis that, by imagining a phenomenon as being different from how it currently is, leads to the isolation of its essential features or aspects, in the sense that such features or aspects cannot be varied or deleted without preventing the phenomenon from being the kind of phenomenon that it is (Parnas and Zahavi, 2002). Another example of a so-called applied phenomenological method is "microphenomenology" (Petitmengin et al., 2018; Depraz, 2020). This method, like Giorgi's, also diverges from the original Husserlian philosophical method. In addition to the method of analysis, micro-phenomenology includes some "principles" regarding the interview. Microphenomenological analysis seeks to identify generic pre-reflexive structures from descriptions of "singular" lived experiences. The pre-reflexive aspect of experience is conceived as experientially "unnoticed," in the sense that it is not immediately accessible to reflective consciousness and verbal description. However, at least in the way Petitmengin et al. (2018) present it, what results from the analysis seems to be more a description of the figurative aspects or features of the object rather than experiential structures of the subject (for example, size, shape, temperature, color, etc.,).

DISCUSSION

Whether to find evidence supporting already-existing insights and concepts or to make it possible for new insights and concepts to emerge from the data itself, phenomenological empirical research must take on board patients' accounts of their subjective experience. Phenomenological interviews should present clear guidelines on both how to conduct them and the qualitative methods employed in analyzing patients' subjective experiences. The research report should follow standards for presenting qualitative research (O'Brien et al., 2014). Still, the most challenging aspect of phenomenological empirical research in psychiatry and psychology is the proper method for analyzing patients' reports. Neither the original Husserlian question of phenomenological philosophizing nor the phenomenological method of philosophical analysis appears appropriate for empirical application. There seems to be a gap between the phenomenological philosophical method and its empirical versions.

Phenomenological philosophy, psychiatry, and psychology have different aims and practical implications. This implies that the methods used in each of these research fields are necessarily different, since they serve as a means to achieve the different aims pursued by each of the corresponding disciplines. In philosophy, the phenomenological method serves as a means to reflect upon fundamental ontological questions regarding our active subjective involvement in the constitution of the world. However, in phenomenological psychiatry and psychology, the methods serve as a means to achieve more precise, complete, and differential diagnoses, with the aim of improving psychotherapy and, ultimately, patients' well-being. Nevertheless, regardless of their divergence from the original philosophical method, Georgi's method of content analysis (to a greater extent), and "microphenomenology" (to a lesser extent), have been quite influential, precisely because of their attempt to bridge this gap, providing a response to the need for a phenomenological method for qualitative research.

An entirely different way of dealing with this problem would not be to seek empirical adaptations of the original phenomenological method inherent in philosophy, nor to limit phenomenology to a mere descriptive task of subjective experience, but to make phenomenology a theoretical framework for empirical research, and even more, a transcendental paradigm. Although its method is certainly fundamental to it, phenomenology should not be reduced to its methodology. Phenomenology is a comprehensive theoretical framework that has been developed on the basis of serious conceptual and empirical research into the subject-world correlation (Zahavi, 2019a), including studies of formal structures of experience (spatiality, temporality, corporeality, intersubjectivity, and historicity), research into the modes of intentionality (perception, agency, phantasy, memory, emotions, and empathy), and psychological analyses of meaning-making processes in social interactions. Additionally, despite the different aims and methods involved, just as in phenomenological philosophy, in phenomenological psychiatry and psychology the core philosophical commitment regarding a critical stance on the conception of reality as mind-independently given is fundamental (Zahavi, 2017, 2019e). Does psychiatry and psychology really need the Husserlian method to adopt the phenomenological attitude toward the conception of reality as mind-dependent? No, because this core philosophical commitment already constitutes the basis of a transcendental paradigm in phenomenological psychiatry and psychology.

Mainstream psychiatry has been developed within a natural-scientific paradigm. From the positivist viewpoint of psychiatry, the notion of normality is defined with regard to the degree of correspondence between subjective experience and objective reality.

Consequently, abnormality is defined in terms of its degree of deviation from an objective reality that provides the evidence for commonsensical validity. For its part, phenomenological psychopathology approaches mental phenomena in terms of a phenomenological analysis of the patient's subjectivity, placing the focus on the conditions of possibility of human experience in general, beyond it being diagnosed as abnormal according to common standards of objectivity. For instance, in current diagnostic systems, psychosis is diagnosed by the presence of hallucinations and delusions, as defined by a "natural attitude" that takes for granted the validity of an objective given reality. In DSM-5, hallucination is defined as a perception without object (or an error of perception) and delusion as a false belief of reality (American Psychiatric Association, 2013). In contrast, from a phenomenological approach, a disturbance is not approached in terms of the clinician's evidence of the inexistence of the object of perception or the lack of external evidence of the patient's belief, but rather in terms of an analysis of the particular mode of intentionality that constitutes the hallucination or delusion as such. In other words, the clinician is concerned with a phenomenological analysis of the patient's subjectivity, addressing with empathic understanding the patient's "self-evidence" or "solipsistic truth," correlated with the experience of hallucination or delusion, respectively. Indeed, the "external" inexistent object should provide for the clinician with evidence that hallucination is not perception, as it is impossible to have a perception without a directly present object.

Consequently, it would be misleading to conceive of hallucination as something to do with perception at all. Instead, hallucinations would have more to do with the phenomenology of fantasy, whose distinctive character is to "re-present" an object of perception that is not directly present, but absent from the actual field of perception. According to Cavallaro (2017), it is not the presentation/re-presentation dichotomy, but what Husserl calls "ego-splitting" (*Ichspaltung*) that is crucial to distinguishing when experiencing the "quasi perception" produced by fantasy and not a perception as such. Ego-splitting makes possible the experience of the "as if" fictive character of self-awareness when fantasizing. However, when hallucinating, the patient experiences his/her own thoughts, anticipations, or imaginations just as in original experiences of perception. So, it may be posited that it is precisely this lack of the "as if" self-awareness of the "quasi perception" that lies at the core of psychosis. Such a theory would require further phenomenological research to draw more distinctions between the nature of hallucination in contrast to that of fantasy, as well as regarding other modalities of experiencing which do not have an intentional object directly present, such as anticipations, thoughts, memories, and dreams. Still, introducing the concept of "ego-splitting" as non-pathological might be challenging to traditional psychiatric concepts, especially with regard to schizophrenia.

Finally, the phenomenological attitude should not be conceived of as being like any other attitude; it is obviously not literally an attitude. The phenomenological attitude is a paradigmatic commitment of a non-pregiven reality. This core philosophical commitment is particularly important because it entails a quite unique approach to mental illness, including different conceptualizations of psychopathology, diagnosis, normality, empathy, and psychotherapy, thus leading qualitative empirical research in psychiatry and psychology toward new horizons. Moreover, the notion of suspending the natural attitude to approaching reality (including all kinds of phenomena) lies at the heart of the phenomenological framework for anyone claiming to be a phenomenologist, whether conceptual or empirical, and regardless of other particular methods and topics of study. In this way, the phenomenological attitude might be conceived of the basis of a transcendental scientific paradigm for qualitative research in psychiatry and psychology. This latter claim, which supports the idea that phenomenological psychology – in order to be properly phenomenological – must become transcendental, and the phenomenological conceptualization of hallucination as pathology of fantasy provide challenging directions for future research.

AUTHOR CONTRIBUTIONS

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

FUNDING

This article has been financially supported by Chilean National Commission for Scientific and Technological Research CONICYT PFCHA/POSTDOCTORADO EN EL EXTRANJERO BECAS CHILE/2017 – 74180011.

REFERENCES

American Psychiatric Association (2013). *Diagnostic and statistical manual of mental disorders DSM-5. 5th Edn.* Washington: American Psychiatric Publishing.

Barbour, R. S. (2000). The role of qualitative research in broadening the "evidence base" for clinical practice. *J. Eval. Clin. Pract.* 6, 155–163. doi: 10.1046/j.1365-2753.2000.00213.x

Barbour, R., and Barbour, M. (2003). Evaluating and synthesizing qualitative research: the need to develop a distinctive approach. *J. Eval. Clin. Pract.* 9, 179–186. doi: 10.1046/j.1365-2753.2003.00371.x

Bornat, J. (2008). "Biographical methods" in *The sage handbook of social research methods*. eds. P. Alasuutari, L. Bickman and J. Brannen (London: Sage), 344–356.

Campbell, L. F., Norcross, J. C., Vasquez, M. J. T., and Kaslow, N. J. (2013). Recognition of psychotherapy effectiveness: the APA resolution. *Psychotherapy* 50, 98–101. doi: 10.1037/a0031817

Cavallaro, M. (2017). The phenomenon of ego-splitting in Husserl's phenomenology of pure phantasy. *J. Br. Soc. Phenomenol.* 48, 162–177. doi: 10.1080/00071773.2016.1250436

Chaminade, T., and Decety, J. (2002). Leader or follower? Involvement of the inferior parietal lobule in agency. *Neuroreport* 13, 1975–1978. doi: 10.1097/00001756-200210280-00029

Creswell, J. W., and Poth, C. N. (2018). *Qualitative inquiry and research design: Choosing among five approaches. 4th Edn.* London: Sage.

Deacon, B. (2013). The biomedical model of mental disorder: a critical analysis of its validity, utility, and effects on psychotherapy research. *Clin. Psychol. Rev.* 33, 846–861. doi: 10.1016/j.cpr.2012.09.007

Deacon, B., and McKay, D. (2015). The biomedical model of psychological problems: a call for critical dialogue. *Behav. Ther.* 38, 231–235.

Depraz, N. (2020). "Husserlian phenomenology in the light of microphenomenology" in *Husserl, Kant and transcendental phenomenology*. eds. I. Apostolu and C. Serban (Berlin: De Gruyter).

Depraz, N., and Desmidt, T. (2019). Cardiophenomenology: a refinement of neurophenomenology. *Phenomenol. Cogn. Sci.* 18, 493–507. doi: 10.1007/s11097-018-9590-y

Depraz, N., Gyemant, M., and Desmidt, T. (2017). A first-person analysis using third-person data as a generative method a case study of surprise in depression. *Constr. Found.* 12, 190–203. Available at: http://constructivist.info/12/2/190

Depraz, N., Varela, F. J., and Vermersch, P. (2003). *On becoming aware: A pragmatics of experiencing*. (Amsterdam: John Benjamin).

Englebert, J., Monville, F., Valentiny, C., Mossay, F., Pienkos, E., and Sass, L. (2019). Anomalous experience of self and world: administration of the EASE and EAWE scales to four subjects with schizophrenia. *Psychopathology* 52, 294–303. doi: 10.1159/000503117

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

Fischer, C. T. (2006). *Qualitative research methods for psychologists: Introduction through empirical case studies*. San Diego, CA: Academic Press.

Fischer, C. T. (2009). Bracketing in qualitative research: conceptual and practical matters. *Psychother. Res.* 19, 583–590. doi: 10.1080/10503300902798375

Fuchs, T. (2005). Corporealized and disembodied minds: a phenomenological view of the body in melancholia and schizophrenia. *Philos. Psychiatry Psychol.* 12, 95–107. doi: 10.1353/ppp.2005.0040

Fuchs, T. (2010a). Subjectivity and intersubjectivity in psychiatric diagnosis. *Psychopathology* 43, 268–274. doi: 10.1159/000315126

Fuchs, T. (2010b). The psychopathology of hyperreflexivity. *J. Specul. Philos.* 24, 239–255. doi: 10.1353/jsp.2010.0010

Fuchs, T. (2013a). "The self in schizophrenia: Jaspers, Schneider and beyond" in *One century of Karl Jaspers' general psychopathology*. eds. G. Stanghellini and T. Fuchs (Oxford: Oxford University Press), 245–257.

Fuchs, T. (2013b). Existential vulnerability: toward a psychopathology of limit situations. *Psychopathology* 46, 1–8. doi: 10.1159/000351838

Fuchs, T., and Koch, S. C. (2014). Embodied affectivity: on moving and being moved. *Front. Psychol.* 5:508. doi: 10.3389/fpsyg.2014.00508

Fuchs, T., Messas, G. P., and Stanghellini, G. (2019). More than just description: phenomenology and psychotherapy. *Psychopathology* 52, 63–66. doi: 10.1159/000502266

Fuchs, T., and Schlimme, J. E. (2009). Embodiment and psychopathology: a phenomenological perspective. *Curr. Opin. Psychiatry* 22, 570–575. doi: 10.1097/YCO.0b013e3283318e5c

Galbusera, L., Fellin, L., and Fuchs, T. (2019). Towards the recovery of a sense of self: an interpretative phenomenological analysis of patients' experience of body-oriented psychotherapy for schizophrenia. *Psychother. Res.* 29, 234–250. doi: 10.1080/10503307.2017.1321805

Galbusera, L., Finn, M. T., and Fuchs, T. (2018). Interactional synchrony and negative symptoms: an outcome study of body-oriented psychotherapy for schizophrenia. *Psychother. Res.* 28, 457–469. doi: 10.1080/10503307.2016.1216624

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

Gallagher, S., and Zahavi, D. (2008). *The phenomenological mind: An introduction to philosophy of mind and cognitive science. 2nd Edn.* New York: Routledge.

Giorgi, A. (2009). *The descriptive phenomenological method in psychology: A modified Husserlian approach*. Pittsburgh, PA: Duquesne University Press.

Giorgi, A. (2012). The descriptive phenomenological psychological method. *J. Phenomenol. Psychol.* 43, 3–12. doi: 10.1163/156916212X632934

Gozé, T., Moskalewicz, M., Schwartz, M. A., Naudin, J., Micoulaud-Franchi, J. A., and Cermolacce, M. (2019). Reassessing "praecox feeling" in diagnostic decision making in schizophrenia: a critical review. *Schizophr. Bull.* 45, 966–970. doi: 10.1093/schbul/sby172

Hammersley, M. (2018). Commentary—on the "indistinguishability thesis": a response to Morgan. *J. Mixed Methods Res.* 12, 256–261. doi: 10.1177/1558689818772764

Høffding, S., and Martiny, K. (2016). Framing a phenomenological interview: what, why and how. *Phenomenol. Cogn. Sci.* 15, 539–564. doi: 10.1007/s11097-015-9433-z

Irarrázaval, L. (2013). Psychotherapeutic implications of self disorders in schizophrenia. *Am. J. Psychother.* 67, 277–292. doi: 10.1176/appi.psychotherapy.2013.67.3.277

Irarrázaval, L. (2015). The lived body in schizophrenia: transition from basic self-disorders to full-blown psychosis. *Front. Psychol.* 6:9. doi: 10.3389/fpsyt.2015.00009

Irarrázaval, L. (2018). Vulnerability in schizophrenia: a phenomenological anthropological approach. *Journal of Intercultural Philosophy* 1, 157–167. doi: 10.11588/icp.2018.1.48070

Irarrázaval, L. (2019). Empathy for the foreign experience: a convergent phenomenological definition. *J. Theor. Phil. Psychol.* doi: 10.1037/teo0000128

Irarrázaval, L., and Sharim, D. (2014). Intersubjectivity in schizophrenia: life story analysis of three cases. *Front. Psychol.* 5:100. doi: 10.3389/fpsyg.2014.00100

Koch, S. C., and Fuchs, T. (2011). Embodied arts therapies. *Arts Psychother.* 38, 276–280. doi: 10.1016/j.aip.2011.08.007

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. U. S. A.* 99, 1586–1591. doi: 10.1073/pnas.032658199

Lysaker, P. H., Dimaggio, G., Hamm, J. A., Leonhardt, B. L., Hochheiser, J., and Lysaker, J. T. (2018a). Disturbances in self-experience in schizophrenia: metacognition and the development of an integrative recovery-oriented individual psychotherapy. *Psychopathology* 52, 135–142. doi: 10.1159/000495297

Lysaker, P. H., Irarrázaval, L., Gagen, E., Armijo, I., Ballerini, M., Mancini, M., et al. (2018b). Metacognition in schizophrenia disorders: comparisons with community controls and bipolar disorder: replication with a Spanish language Chilean sample. *Psychiatry Res.* 267, 528–534. doi: 10.1016/j.psychres.2018.06.049

Martin, L. A., Koch, S. C., Hirjak, D., and Fuchs, T. (2016). Overcoming disembodiment: the effect of movement therapy on negative symptoms in schizophrenia—a multicenter randomized controlled trial. *Front. Psychol.* 7:483. doi: 10.3389/fpsyg.2016.00483

Martiny, K. M. (2017). Varela's radical proposal: how to embody and open up cognitive science. *Constr. Found.* 13, 59–67. Available at: http://constructivist.info/13/1/059

Maxwell, J. (2011). *A realist approach for qualitative research*. Thousand Oaks, CA: Sage.

Maxwell, J. (2012). *Qualitative research design: An interactive approach. 3rd Edn.* London: Sage.

Maxwell, J. A. (2019). Distinguishing between quantitative and qualitative research: a response to Morgan. *J. Mixed Methods Res.* 13, 132–137. doi: 10.1177/1558689819828255

Mishara, A. L. (1994). "A phenomenological critique of commonsensical assumptions of DSM-III-R: The avoidance of the patient's subjectivity" in *Philosophical perspectives on psychiatric diagnostic classification*. eds. J. Z. Sadler, M. A. Schwartz and O. P. Wiggins (Baltimore: Johns Hopkins Series in Psychiatry and Neuroscience), 129–147.

Morgan, D. L. (2018). Living within blurry boundaries: the value of distinguishing between qualitative and quantitative research. *J. Mixed Methods Res.* 12, 268–279. doi: 10.1177/1558689816686433

Morrow, S. L. (2005). Quality and trustworthiness in qualitative research in counseling psychology. *J. Couns. Psychol.* 52, 250–260. doi: 10.1037/0022-0167.52.2.250

Nordgaard, J., Sass, L. A., and Parnas, J. (2013). The psychiatric interview: validity, structure and subjectivity. *Eur. Arch. Psychiatry Clin. Neurosci.* 263, 353–364. doi: 10.1007/s00406-012-0366-z

O'Brien, B. C., Harris, I. B., Beckman, T. J., Reed, D. A., and Cook, D. A. (2014). Standards for reporting qualitative research: a synthesis of recommendations. *Acad. Med.* 89, 1245–1251. doi: 10.1097/ACM.0000000000000388

Pallagrosi, M., and Fonzi, L. (2018). On the concept of praecox feeling. *Psychopathology* 51, 353–361. doi: 10.1159/000494088

Pallagrosi, M., Fonzi, L., Picardo, A., and Biondi, M. (2014). Assessing clinician's subjective experience during the interaction with patients. *Psychopathology* 47, 111–118. doi: 10.1159/000351589

Parnas, J., and Handest, P. (2003). Phenomenology of anomalous self-experience in early schizophrenia. *Compr. Psychiatry* 44, 121–134. doi: 10.1053/comp.2003.50017

Parnas, J., and Henriksen, M. G. (2014). Disordered self in the schizophrenia spectrum: a clinical and research perspective. *Harv. Rev. Psychiatry* 22, 251–265. doi: 10.1097/HRP.0000000000000040

Parnas, J., Moeller, P., Kircher, T., Thalbitzer, J., Jannson, L., Handest, P., et al. (2005). EASE: examination of anomalous self-experience. *Psychopathology* 38, 236–258. doi: 10.1159/000088441

Parnas, J., and Sass, L. (2008). "Varieties of phenomenology: on description, understanding, and explanation in psychiatry" in *Philosophical issues in psychiatry: Explanation, phenomenology and nosology*. eds. K. Kendler and J. Parnas (New York: Johns Hopkins University Press), 239–277.

Parnas, J., and Zahavi, D. (2002). "The role of phenomenology in psychiatric diagnosis and classification" in *Psychiatric diagnosis and classification*. eds. M. Maj, W. Gaebel, J. J. López-Ibor and N. Sartorius (Hoboken, NJ: John Wiley & Sons Inc.), 137–162.

Patton, M. (2015). *Qualitative research and evaluation methods. 4th Edn.* Thousand Oaks, CA: Sage.

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

Raballo, A., Sæbye, D., and Parnas, J. (2009). Looking at the schizophrenia spectrum through the prism of self-disorders: an empirical study. *Schizophr. Bull.* 37, 44–351. doi: 10.1093/schbul/sbp056

Röhricht, F., and Papadopoulos, N. (2010). *A treatment manual: Body oriented psychological therapy for chronic schizophrenia*. London: Newham Centre for Mental Health.

Ruby, P., and Decety, J. (2001). Effect of subjective perspective taking during simulation of action: a PET investigation of agency. *Nat. Neurosci.* 4, 546–550. doi: 10.1038/87510

Sass, L. (1992). *Madness and modernism: Insanity in the light of modern art, literature and thought*. New York: Basic Books.

Sass, L., Borda, J. P., Madeira, L., Pienkos, E., and Nelson, B. (2018). Varieties of self disorder: a bio-pheno-social model of schizophrenia. *Schizophr. Bull.* 44, 720–727. doi: 10.1093/schbul/sby001

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

Sass, L., Pienkos, E., Skodlar, B., Stanghellini, G., Fuchs, T., Parnas, J., et al. (2017). EAWE: examination of anomalous world experience. *Psychopathology* 50, 10–54. doi: 10.1159/000454928

Schwartz, H., and Jacobs, J. (1979). *Qualitative sociology: A method to the madness*. New York: Free Press.

Škodlar, B., and Henriksen, M. H. (2019). Toward a phenomenological psychotherapy for schizophrenia. *Psychopathology* 52, 117–125. doi: 10.1159/000500163

Tashakkori, A., and Teddlie, C. (2010). *SAGE handbook of mixed methods in social & behavioral research. 2nd Edn.* Thousand Oaks, CA: SAGE Publications, Inc.

Varela, F. (1990). *Conocer: Las ciencias cognitivas: tendencias y perspectivas. Cartografía de las ideas actuales [To know: Cognitive sciences: Trends and prospects. Cartography of current ideas]*. Barcelona: Gedisa.

Varela, F. J. (1996). Neurophenomenology: a methodological remedy for the hard problem. *J. Conscious. Stud.* 3, 330–349.

Varela, F. J., and Shear, J. (1999). First-person methodologies: what, why, how? *J. Conscious. Stud.* 6, 1–14.

World Health Organization (2012). *The ICD-10 classification of mental and behavioral disorders: Clinical descriptions and diagnostic guidelines*. Geneva: World Health Organization.

Zahavi, D. (2015). You, me and we: the sharing of emotional experiences. *J. Conscious. Stud.* 22, 84–101.

Zahavi, D. (2017). *Husserl's legacy: Phenomenology, metaphysics, and transcendental philosophy*. Oxford, UK: Oxford University Press.

Zahavi, D., and Martiny, K. M. M. (2019). Phenomenology in nursing studies: new perspectives. *Int. J. Nurs. Stud.* 93, 155–162. doi: 10.1016/j.ijnurstu.2019.02.018

Zahavi, D. (2019a). *Phenomenology: The basics*. London: Routledge.

Zahavi, D. (2019b). Getting it quite wrong: Van Manen and Smith on phenomenology. *Qual. Health Res.* 29, 900–907. doi: 10.1177/1049732318817547

Zahavi, D. (2019c). Applied phenomenology: why it is safe to ignore the epoché. *Cont. Philos. Rev.* 1–15. doi: 10.1007/s11007-019-09463-y

Zahavi, D. (2019d). The practice of phenomenology: the case of Max van Manen. *Nurs. Philos.* 21:e12276. doi: 10.1111/nup.12276

Zahavi, D. (2019e). "Phenomenology as metaphysics" in *The Routledge handbook of metametaphysics*. eds. R. Bliss and J. Miller (London: Routledge).

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

ERRATUM
published: 23 July 2020
doi: 10.3389/fpsyg.2020.01989

Erratum: A Phenomenological Paradigm for Empirical Research in Psychiatry and Psychology: Open Questions

*Frontiers Production Office**

Frontiers Production Office, Frontiers Media SA, Lausanne, Switzerland

Keywords: applied phenomenology, methodology, qualitative research, psychiatry and psychology, phenomenological interviews

An Erratum on

A Phenomenological Paradigm for Empirical Research in Psychiatry and Psychology: Open Questions
by Irarrázaval, L. (2020). Front. Psychol. 11:1399. doi: 10.3389/fpsyg.2020.01399

Due to an error in the typesetting process, in the last paragraph of the Introduction, the term "*epoché*" was erroneously replaced by the term "*what it is like.*" A correction has therefore been made to the **Introduction** section, paragraph four:

"This article will address the "What," "Why," and "How" of phenomenological interviews, reviewing recent empirical research in the field of phenomenological psychopathology and psychotherapy. Important to note is that qualitative research, as described above, refers to empirical research, not to basic or theoretical investigations. Phenomenological qualitative research in psychology has been developed using Husserlian concepts such as the "*epoché*" and the "phenomenological reduction," and precisely on the use of such conceptualizations is where most of the current discussion has been placed. The article, therefore, will not attempt to provide a broad understanding of the phenomenological tradition. Instead, it will focus on a more specific discussion of methodological issues concerning the empirical application of phenomenology in qualitative research in psychiatry and psychology, and Husserl's methodology in particular. To do so, we first need to agree that the application of phenomenology to empirical research in psychiatry and psychology employing interviews is qualitative, not quantitative. In a strict sense, quantitative methodology based on frequency and scales of severity of the patients' anomalous experience, although necessary for the statistical validation of the interviews, goes beyond the scope of phenomenology. According to the phenomenological approach, mental disorders cannot be reducible to a cerebral organic basis, nor to numbers, as they are not entities *per se* but psychopathological configurations that can be identified in the diagnostic process of interaction between a clinician and a patient (Fuchs, 2010; Pallagrosi et al., 2014; Pallagrosi and Fonzi, 2018; Gozé et al., 2019). Consequently, phenomenological interviews are designed to address not objective, but subjective data, namely the *what it is like* of patients' anomalous experiences. In this way, the patients' descriptions of their subjective experiences are not conceived as "static" entities, but, rather, as part of dynamically, open-ended developing processes and interpretations (Martiny, 2017).

The publisher apologizes for this mistake. The original article has been updated.

Approved by:
Frontiers Editorial Office, Frontiers Media SA, Switzerland

***Correspondence:**
Frontiers Production Office
production.office@frontiersin.org

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

Received: *16 July 2020*
Accepted: *17 July 2020*
Published: *23 July 2020*

Citation:
Frontiers Production Office (2020) Erratum: A Phenomenological Paradigm for Empirical Research in Psychiatry and Psychology: Open Questions. Front. Psychol. 11:1989. doi: 10.3389/fpsyg.2020.01989

REFERENCES

Fuchs, T. (2010). Subjectivity and intersubjectivity in psychiatric diagnosis. *Psychopathology* 43, 268–274. doi: 10.1159/000315126

Gozé, T., Moskalewicz, M., Schwartz, M. A., Naudin, J., Micoulaud-Franchi, J. A., and Cermolacce, M. (2019). Reassessing "praecox feeling" in diagnostic decision making in schizophrenia: a critical review. *Schizophr. Bull.* 45, 966–970. doi: 10.1093/schbul/sby172

Martiny, K. M. (2017). Varela's radical proposal: how to embody and open up cognitive science. *Constr. Found.* 13, 59–67. Available at: http://constructivist.info/13/1/059

Pallagrosi, M., and Fonzi, L. (2018). On the concept of praecox feeling. *Psychopathology* 51, 353–361. doi: 10.1159/000494088

Pallagrosi, M., Fonzi, L., Picardo, A., and Biondi, M. (2014). Assessing clinician's subjective experience during the interaction with patients. *Psychopathology* 47, 111–118. doi: 10.1159/000351589

frontiers in Psychology

HYPOTHESIS AND THEORY
published: 14 July 2020
doi: 10.3389/fpsyg.2020.01548

The Living Transcendental — An Integrationist View of Naturalized Phenomenology

*Thomas Netland**

Department of Philosophy and Religious Studies, Norwegian University of Science and Technology, Trondheim, Norway

In this article I take on the "Transcendentalist Challenge" to naturalized phenomenology, highlighting how the ontological and methodological commitments of Merleau-Ponty's philosophy point in the direction of an integration of the transcendental and the scientific, thus making room for a productive exchange between philosophy and psychological science when it comes to understanding consciousness and its place in nature. Discussing various conceptions of naturalized phenomenology, I argue that what I call an "Integrationist View" is required if we are to make sense of the possibility of productive exchange between phenomenology and the sciences. My main argument is that if we conceive of consciousness as a structure of behavior ontologically prior to the distinctions between objectivity and subjectivity and third- and first-person perspectives, we arrive at a view of the transcendental as not essentially separate from the domain of science, but rather as contingent organizational norms of empirical nature that are best illuminated through a dialectical exchange between phenomenological and scientific approaches. I end by showing how Merleau-Ponty's engagement with the "Schneider case" in an example of such an integration.

Keywords: naturalized phenomenology, Merleau-Ponty, transcendental philosophy, Schneider, enactivism

OPEN ACCESS

Edited by:
Wenjing Cai,
Shanghai Jiao Tong University, China

Reviewed by:
Saulius Geniusas,
The Chinese University of Hong Kong, China
Komarine Romdenh-Romluc,
University of Sheffield, United Kingdom

***Correspondence:**
Thomas Netland
thomas.netland@ntnu.no

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

Received: *11 March 2020*
Accepted: *09 June 2020*
Published: *14 July 2020*

Citation:
Netland T (2020) The Living Transcendental — An Integrationist View of Naturalized Phenomenology. Front. Psychol. 11:1548. doi: 10.3389/fpsyg.2020.01548

INTRODUCTION

During the last decades, phenomenology has become increasingly influential as a resource for developments in the mind sciences. This is especially so in the research program known as "the enactive approach," one of the central tenets of which is, in the words of its cofounder Evan Thompson, that "[i]t is not only possible, but also necessary, to pursue phenomenology and experimental science as mutually constraining and enlightening projects" (2007, p. 273). The prospects of such a relationship is, however, not without difficulties, but has been challenged both by people skeptical of phenomenology's credentials altogether and by phenomenologists who reject the idea of "naturalizing" a philosophy that, in their view, is concerned with the conditions that *enable* scientific thinking in the first place and as such cannot be informed by its results. This latter, "Transcendentalist Challenge" to naturalized phenomenology, is the motivating force for this paper.

My overarching aim in what follows is to propose the position I call the "Integrationist View" (IV), which consists in a reconceptualization of the notions of the "transcendental" and "nature" in a way that allows for a methodological and ontological integration of scientific and phenomenological perspectives. In outlining this view, I draw on the early works of Maurice Merleau-Ponty, the classical phenomenologist known for his extensive engagement with scientific literature. I am far from the first to argue that Merleau-Ponty's philosophy is a promising starting

point for making sense of the project of naturalizing phenomenology. At the same time, as Jack Reynolds has recently observed, "exactly how to understand the Gordian knot concerning Merleau-Ponty's implicit and explicit commitments regarding transcendental reasoning, phenomenology and empirical science, remains contested, more than 50 years after his death" (2017, p. 85). Thus, although I do not presume to completely resolve this knot here, this paper is also a contribution to discussions in Merleau-Ponty scholarship. The reading I propose emphasizes the significance of his first book, *The Structure of Behavior* (1942/1963; henceforth *Structure*), as a background for making sense of the further development of his thought. As such, my reading is at least partly aligned with and indebted to Toadvine (2009) and Morris (2018), both of which, notwithstanding some interpretative differences, see Merleau-Ponty primarily as a philosopher of nature, one of the key concerns of which was to establish the idea of an immanent, expressive *sense* of nature in the form of the embodied and active structure of living organisms' existence. In this way, I see this paper as a contribution to the project of construing enactivism as a philosophy of nature (e.g., Gallagher, 2017, pp. 21–24).

The crux of my argument is that if we conceive of consciousness as a structure of behavior ontologically prior to the distinctions between objectivity and subjectivity and third- and first-person perspectives, we arrive at a view of the transcendental as *not* essentially separate from the scientific, but rather as contingent organizational norms of nature that are best illuminated through a dialectical exchange between phenomenological and scientific approaches. I start by sketching the general contours of transcendental philosophy and the "Transcendentalist Challenge" to naturalized phenomenology, taking Gardner's (2015) transcendentalist reading of Merleau-Ponty as the point of departure (1). I then turn to Zahavi's (2017) suggestion of two alternative ways to understand what a "naturalized phenomenology" amounts to, arguing that the position I label "Modest Transcendentalism" lacks the resources for making adequate sense of the possibility of a productive exchange between phenomenology and science, and propose that this task rather requires the "Integrationist View" (2). Thereafter, I show how the notion of structures of behavior is apt to yield an integrationist ontology (3) before I return to criticize Gardner's transcendentalist reading of Merleau-Ponty in the context of the phenomenological method (4). Lastly, I propose a way to read Merleau-Ponty's engagement with the "Schneider case" in *Phenomenology of Perception* (1945/2012; henceforth *Phenomenology*) as an instance of the IV in action (5).

THE TRANSCENDENTALIST CHALLENGE AND VARIETIES OF TRANSCENDENTALISM

In "Merleau-Ponty's Transcendental Theory of Perception," (2015) Gardner gives expression to one of the main theoretical challenges to the idea of a naturalized phenomenology, namely, the argument that phenomenology is essentially a form of *transcendental* philosophy and, as such, operates in a domain strictly independent from the scientific.[1] Indeed, the main target of Gardner's paper is what he calls the "Psychological Interpretation," which reads Merleau-Ponty's *Phenomenology* as offering insights about perception that can both be put to use by and find support in empirical cognitive science.[2] On Gardner's reading, on the contrary, Merleau-Ponty's arguments have the same "form and idealistic trajectory" as Immanuel Kant's transcendental philosophy (2015, p. 313) and "involves no positive estimate of psychological science as an independent source of knowledge that philosophy ought to accommodate" (2015, p. 319), leading him to conclude that the naturalistic philosophy of psychology that some find in the *Phenomenology* "has only an oblique relation to the position Merleau-Ponty is actually arguing for" (2015, p. 321). Before looking closer at Merleau-Ponty's position, let us have a look at transcendental philosophy more generally.

What is transcendental philosophy? The history of this notion and the discussions surrounding it shows that it is difficult, if not impossible, to give one precise characterization that covers all its appearances.[3] I will try to give a sense of this fluidity of the notion in what follows, but take this as a first, provisional definition: Transcendental philosophy aims to uncover the ground for objective knowledge, where "ground" is understood not as a Cartesian foundational proposition that secures the possibility of knowledge, but rather as the structures of consciousness constitutive of our knowing. The prime example here is Kant. In *Critique of Pure Reason* (2007), he asked how it is possible for experience to be a source for knowledge and answered that the necessary conditions for this is that experience be oriented in space and time (the forms of intuition), structured in conformity with the categories of the understanding (e.g., causality and substantiality), and unified in relation to the unity of the transcendental subject (the "I think").[4] Notice how this project is fundamentally different from what we find in the sciences. After all, science takes the possibility and validity of experience, objectivity, and knowledge for granted, depending on these in its project of gathering facts and constructing theories about the world. Transcendental philosophy, on the other hand, does not seek fact or theory in the same sense but rather the conditions that make them possible. We can thus see how the idea of a productive exchange between these two domains is problematic: science does not seem to require an understanding of its transcendental conditions of possibility in order to succeed, and transcendental philosophy cannot rely

[1] I'm I am here using Gardner as representative for a concern raised by multiple phenomenologists, e.g., (De Preester, 2002; Moran, 2013).

[2] Gardner's examples include reading Merleau-Ponty as providing "a convincing critique of the representationalism which holds sway in cognitive science" and an "account of skill acquisition [that] stands in deep accord with developments in brain science neural network theory" (2015, p. 297).

[3] Habermas (1991) provides a clear overview of the development and internal and external critiques of the tradition of transcendental philosophy.

[4] This is intended to give a rough grasp—and not in any way to be an exact rendering—of what happens in Kant's first *Critique*.

on scientific findings without presupposing what it seeks to understand.[5]

This distinction is underlined by the fact that transcendental philosophy is a non-empirical, *a priori* endeavor. Consider how Kant deals with the concept of causality: On his view, we do not acquire this concept through experience; rather, it belongs to the subject as one of its necessary *conditions* for experience to be possible in the first place. Methodologically, this means that the transcendental here is identifiable by purely *a priori* means. Exactly how to understand the nature of Kant's transcendental arguments is a discussion in its own right and not something I will dig into here.[6] For our purposes, it suffices to draw attention to one way we can understand the contrast and continuity between Kant and the phenomenological tradition when it comes to the notion of the transcendental. In this context, we can distinguish between two forms of transcendental argument found in Kant—one that is dismissed by the phenomenologists and one that they to some degree take up and refine. In the former, we find progressive arguments aimed at establishing the necessary objective validity of certain concepts (e.g., causality). These lead Kant to construe the transcendental as structures belonging to subjectivity (more precisely to the understanding) independently of any particular experience, which determines in advance the possible form of all future experience. The latter form of argument is regressive, beginning from given facts or experiences and proceeding to reconstruct the conditions for the possibility of their givenness as such.[7] This means that one here is paying more attention to concrete matters and how these are experienced and apprehended compared to in the former case, where the aim rather is to establish the necessary forms all such matters must conform to.

By rejecting the first kind of argument and modifying the second, Edmund Husserl's transcendental phenomenology represents a shift in the notion of the transcendental. Now, the transcendental is conceived not as belonging solely to the subject but to the subject-world *correlation*, not as principles abstractly outlining the form of all possible experience but as structures constitutive of and originating within actual experience.[8] This makes the necessity and a prioricity of the transcendental in Husserl's philosophy quite different from that found in Kant. As Julia Jansen observes (presumably thinking primarily of Kant's progressive arguments),

> Kant thinks of a necessary unity as a unity that receives its necessity "top–down" from the "highest point" of reason [. . .]. Husserl, on the contrary, thinks of unity "laterally," as a unity of "coincidence (*Deckung*)," which enables *a priori* insight not only into necessities that "reason itself produces according to its own plan" (B xiii), as Kant famously claimed, but also into necessities reason genuinely *discovers* (Jansen, 2015, pp. 48–49, emphases in original).

In other words, the transcendental is now understood as in a certain sense experientially discoverable, drawing it closer to the empirical domain. This is evident in Husserl's claim that the proper method of transcendental philosophy should be *description* rather than deduction. On his view, this shift represents a necessary correction of Kant's project, which from the phenomenological perspective takes the form of problematic metaphysics, resting for instance on a misguided separation between sensibility and the understanding (ibid., p. 59). Rather than assuming such a separation and then attempting to identify the contribution of each faculty through a rational construction, phenomenology takes the actuality of perception as its point of departure and seeks to describe, clarify, and analyze the emergence of meaning and objectivity as evident *therein*.

While this surely moves the transcendental domain closer to the empirical relative to what we find in Kant, it does not entail that the distinction between the transcendental and scientific domains collapses. In Husserl's phenomenology, the key methodological tools for arriving at the domain proper to transcendental phenomenology are the *epoché* and the *reduction*. The *epoché* amounts to a shift from the "natural" to the "phenomenological" attitude through *bracketing* or *suspending* our normal interest in and presuppositions regarding the external world as such, so as to focus on the subject-world correlation—i.e., on the *how* of experience rather than the *what* of the experience*d*. The reduction is then the next step, consisting in the systematic examination of this correlational structure in light of its transcendental function.[9] This now marks the difference between phenomenology (*qua* transcendental) and science. While they both might take their data from experience, their attitudes are fundamentally distinct—the latter seeks to know the *objects* of experience and takes their existence as such for granted, whereas the former aims to clarify the constitutive structures of the givenness of the world thanks to which it appears as objective, meaningful, etc. In *Rethinking Transcendentalism: the Limits of Transcendental Reflection*, we will look closer at Merleau-Ponty's verdict of this method, which famously is that "the most important lesson of the reduction is the impossibility of a complete reduction" (2012, p. lxxvii).

The transcendental conditions identified through the phenomenological method are of a quite different sort from those deduced by Kant. Here, we move from the form-imposing role of the categories of the understanding, to constitutive conditions "visible" within experience, such as consciousness' horizonal structures (its co-intention of "absent" and indeterminate features such as past and future, the hidden profiles of visual objects, etc.). Again, this means that we are operating with a quite different notion of "transcendental" here than what we started out with. As Jansen suggests, Kant would probably have dismissed the transcendental structures identified by Husserl

[5]This is not to say that transcendental philosophy has no interest in science. On the contrary, its main motivation has traditionally been to establish a philosophical ground for scientific knowledge.

[6]But see, e.g., Strawson (1966); Henrich (1969), Stern (2000), and Ameriks (2003).

[7]For instance, what structures are constitutive of operations in pure arithmetic? (Notice that the question here is neither *whether* it is possible nor whether its results will be universally true). Kant's answer is that this possibility, among other things, depends on time as a pure, subjective form of intuition (2007, B17, B56).

[8]Zahavi (1996) argues convincingly that another way in which Husserl transforms Kant's transcendental philosophy, is by recognizing the constitutive role of intersubjectivity, thus going beyond the emphasis on the individual consciousness.

[9]There are competing interpretations concerning the exact nature of the *epoché* and the reduction in Husserl's phenomenology. Here, I'm I am relying on Dan Zahavi's illuminative rendering in *Husserl's Legacy* (2017, pp. 56–60).

as "crude empirical generalities" (2015, p. 78) that fails to meet his strict criteria for *a priori* necessity. As we will see in the next section, some of these "generalities" identified by phenomenology—more precisely, the invariances of experience disclosed by eidetic analyses[10]—seems to mark a point of contact between phenomenology and psychological science on Husserl's view. A discussion of what this means for the prospects of a naturalized phenomenology will have to wait until then.

For now, these are the key takeaways. Despite his departure from Kant's method and metaphysics, Husserl maintains the distinction between transcendental and scientific enterprises. However, the transcendental is now understood as "closer" to the empirical, giving a new sense to its necessary and *a priori* status (invariant/essential constitutive structures *of* experience rather than forms logically imposed upon it). The distinction between science and transcendental philosophy is maintained but now understood as one of *attitudes*. The prospects for a mutually informative relation between them still looks dim—after all, one presupposes the attitude which the other suspends and analyses, and more generally, they are simply preoccupied with different kinds of questions.

Let us now return to Gardner's transcendentalist interpretation of Merleau-Ponty's *Phenomenology*. On his view, transcendental philosophy here undergoes yet a transformation. That is, he sees Merleau-Ponty as establishing *preobjective* perception as "a ground-level transcendental condition" (2015, p. 307). As such, Merleau-Ponty's philosophy entails a critique of both Kant and Husserl's transcendental projects to the extent that they are characterized by "objective thought." "Objective thought" here means a certain dogmatic way of accounting for experience's "articulation into objects and its character [. . .] as involving a relation of subject to object" (ibid, p. 301). The intellectualist tendencies of Merleau-Ponty's transcendentalist predecessors fall into this category due to their taking thoughts about objects as the ultimate *explanans* (i.e., objective thought is responsible for the objectual character of experience). A different form of the same dogma is shared by the view we can call scientific realism or naturalism ("empiricism" in Merleau-Ponty's terminology), which takes the objectual character of experience to be caused by a subject-independent world already articulated into objects. In short, both intellectualism and empiricism take objectivity as a *given*, and Merleau-Ponty's transcendental project consists in disclosing the *origin* of objective thought as such from the preobjective and ambiguous perceptual field. It is on this basis that Gardner dismisses interpretations of Merleau-Ponty that see him as providing, among other things, "arguments for the dependence [. . .] of consciousness on embodiment [and] a convincing critique of the representationalism which holds sway in cognitive science" (2015, p. 297). Such psychological readings take consciousness and perception as *objects* to be described and explained, rather than as the field where objectivity emerges in the first place, and thus leaps over the issue that actually drives the *Phenomenology*.

Although I disagree with the strict separation Gardner sets up between transcendentalism and science, I think his claim that Merleau-Ponty's transcendental philosophy is *ontological* in form is basically correct.[11] As he says, "talk of pre-objective being is not just talk of *experience* prior to the involvement of objectivity concepts in experience: it is talk of experienced *being* which is pre-objective" (2015, p. 298, emphases in original). The *Phenomenology*'s critique of "objective thought," then, is not merely a critique of certain conceptions of *experience*; rather, it is "a critique of the *metaphysical* claim that objective representation is adequate to the representation of reality or, put the other way around, that reality is as objectivity concepts represent it as being" (ibid.). As we will see later, this ontological dimension of Merleau-Ponty's project is a key element in the Integrationist View I am suggesting. In other words, my point of divergence from Gardner's interpretation concerns what the metaphysics of preobjectivity entails for the prospects of a mutually enlightening relation between phenomenology and science. On my view, which conceives it as a continuation of the project *Structure* sets in motion, Merleau-Ponty's ontology represents a promising step toward a "phenomenologizing" of nature where the border between the transcendental and the scientific becomes diffused. For Gardner, on the contrary, the transcendental nature of Merleau-Ponty's project means that its extensive engagement with scientific literature must be understood merely as, in Reynolds' apt words, a sort of "Wittgenstein's ladder," which should be kicked away once the transcendental is reached (2017, p. 98). "Engagement with scientific psychology," Gardner claims,

> sharpens and refines our appreciation of psychological considerations, which in turn helps us to reach a position from which phenomenological truth can be grasped on the basis of an apodictic relation to the pre-objective, rendering transcendental reflection strictly independent of any application of the scientific method (Gardner, 2015, p. 319).

The idea here is that considerations of psychological science might serve the instrumental role of ridding transcendental philosophy of its intellectualist pretensions, but that transcendental reflection proper gets underway only after this labor is done with and then within an autonomous domain indifferent to the scientific. On this point, then, Merleau-Ponty appears to be fully in line with his transcendentalist predecessors. "Merleau-Ponty," Gardner says, "provides [. . .] many statements of how the conditions that his phenomenology uncovers are intended to be in the true and genuine sense transcendental, i.e., *a priori* and necessary, and non-identical with empirical, contingent, or mundane states of affairs" (2015, p. 300). A legitimate question here, however, is how it is possible for transcendental philosophy to be both *reformed* by (ibid, p. 319) *and* "strictly independent of" considerations from scientific psychology. Can one really have both?

[10]The method of eidetic analysis consists, in short, in an imaginative variation of the phenomenon in question in order to become aware of its invariant structure or *essence*—that without which it would no longer be what it is. In this way, it aims to answer questions such as "what do essentially characterize acts of perceiving, imagining, remembering, judging, etc., and how are these different acts related to each other?" (Zahavi, 2017, p. 15).

[11]I do, however, not agree that Merleau-Ponty's transcendentalism warrants the label of an "idealist metaphysics" (2015, pp. 309). I'll I will not be able to explicitly argue this point here, but I take my case for an Integrationist View of naturalized phenomenology to provide some reasons for rejecting such a description.

Gardner's insistence on a strict independence of the transcendental from the scientific seems to stem from a specific conception of what a mutually enlightening relation between them would have to amount to. According to him, naturalized interpretations of Merleau-Ponty's project will not only have the consequence that his philosophy can "become subject [. . .] to empirical correction" but also that "the task of explanation [. . .] tends inevitably to pass out of the hands of phenomenology into neurophysiology and other more empirically tough-minded quarters" (2015, p. 297). In other words, it would ultimately amount to an unequivocal abandonment of transcendental philosophy in favor of an all-encompassing scientific naturalism, where all legitimate questions are seen as answerable by the methods of natural science. There surely are those who advocate this form of "naturalized phenomenology."[12] We find a prime example in the introduction to the anthology *Naturalized Phenomenology*, where the editors explicitly state that "naturalized" here means "integrated into an explanatory framework where every acceptable property is made continuous with the properties admitted by the natural sciences" (Roy et al., 1999, pp. 1–2). Now, if this is what one means by naturalization, and one by "phenomenology" refers to the philosophical tradition of Husserl and Merleau-Ponty, then the notion of a naturalized phenomenology is an impossibility—a category mistake (Zahavi, 2013, p. 30). After all, phenomenology thus conceived would have to partake without question in the "natural attitude" or "objective thought," giving up on its defining task of clarifying the constitution and/or origin of objectivity and thus ceasing to be phenomenology altogether.

That being said, the choice Gardner seems to presuppose between, on the one hand, a transcendental philosophy indifferent to the results of science and, on the other, a naturalistic philosophy that yields all authority to such findings is a false dichotomy. As Zahavi (2013, 2017) suggests, there seems to be at least two alternative conceptions of what a naturalized phenomenology can amount to available. Below I present these, arguing that we should favor the alternative I label the Integrationist View over the more conservative Modest Transcendentalism before I, in the remainder of this text, propose a reading of Merleau-Ponty's philosophy as an instance of the former.

FROM MODEST TRANSCENDENTALISM TO THE INTEGRATIONIST VIEW

While Zahavi on multiple occasions has raised concerns over "naturalizing" approaches to phenomenology from a transcendentalist perspective (2004; 2010; 2011; 2013; and 2017), he is simultaneously one of our contemporary phenomenological philosophers that has done most to facilitate and engage in fruitful dialogs with psychological science. It is thus not surprising that we in his writings find suggestions for conceptions of naturalized phenomenology where the philosophical or transcendental core of phenomenology is maintained. Zahavi tends to point to two such alternatives. The first keeps the idea of phenomenology as transcendental philosophy where "transcendental" entails belonging to a domain strictly separate from the natural and scientific but is more liberal than Gardner's view in that it, nonetheless, allows for some form of mutual enlightenment between the two domains (2017, pp. 162–163). The second alternative is based on rethinking the very notions of the "transcendental" and the "natural" as traditionally conceived, pushing for a tighter integration of phenomenology and science within the framework of a "phenomenologized nature" (2017, p. 167). I call these Modest Transcendentalism (MT) and the Integrationist View (IV), respectively.

Zahavi makes it clear that he is sympathetic to both alternatives and emphasizes that they "should not be seen as incompatible alternatives between which we have to choose" but that "they might be pursued simultaneously" (2017, p. 169). He argues convincingly that Husserl, despite his antinaturalist reputation, subscribed to MT and suggests that he might even have accepted the more radical IV (2017, p. 168). Thus, although I, in the following, use Zahavi's reading of Husserl as representative of MT, I do not assume either of them to be unequivocally committed to this view. I will, however, dispute Zahavi at one account: his compatibility claim quoted above, which he makes without elaboration. How can the two alternatives be compatible? After all, IV aims to unsettle a core pillar of MT's framework, namely, the separation of the transcendental from the natural. As long as this is what defines the difference between the two alternatives, it seems that we *do* have to choose between them. If that is right, I believe that IV has the stronger case. The reason for this is that, when pressed to make adequate sense of the relationship it sets up between phenomenology and science, MT seems to have difficulties preserving the traditional transcendental–natural distinction it presumably subscribes to and to inadvertently and implicitly collapse into a view more like IV. In other words, my argument in what follows is that IV is best suited to give weight to and make coherent the productive exchange between philosophical and scientific perspectives envisioned by MT.

What does the exchange between phenomenology and science consist in for MT? Zahavi suggests the following:

> Phenomenology can question and elucidate basic theoretical assumptions made by empirical science, just as it might aid in the development of new experimental paradigms. Empirical science can present phenomenology with concrete findings that it cannot simply ignore, but must be able to accommodate (2017, p. 162).

Through its eidetic analyses of consciousness, phenomenology yields descriptions and theories of phenomena such as perception, imagination, embodiment, etc., which can serve as basis for engaging critically with scientists' assumptions regarding the same phenomena.[13] Notice that we, in the quote's second sentence, find a clear contrast to the view

[12] Although I doubt Gardner's (2015, p. 297) prime examples of psychological interpreters of Merleau-Ponty—Hubert Dreyfus, Shaun Gallagher, and Sean Kelly—would fit easily into that category.

[13] Gallagher's (2003) idea of "front-loading" phenomenology is a good illustration of how phenomenology can also be used to inform experimental settings.

offered by Gardner's Transcendental Interpretation of Merleau-Ponty, which, remember, rejects the idea of science as "an independent source of knowledge that philosophy ought to accommodate" (2015, p. 319). The question now is, how does this relationship suggested by MT square with the idea that phenomenology and science belong to two essentially separate domains? That is, *in virtue of what* is phenomenology justified as having a say concerning scientific theory, and empirical findings an impact on phenomenology?

MT's first response is that we here are not yet talking about *transcendental* phenomenology. Zahavi reminds us of Husserl's view that "to engage in an eidetic and *a priori* analysis of experiential consciousness is to do psychology—and not yet phenomenology proper" (2017, p. 157). MT's commitment to a separation of the transcendental and the scientific together with its opening for a relationship of mutual enlightenment between phenomenology and science thus seems to rest upon a distinction between two forms of phenomenology—transcendental and psychological. Here, the latter is understood as remaining within the natural attitude, studying consciousness for its own sake in a non-reductive way, whereas the former is interested in consciousness "insofar as [it] is taken to be a condition of possibility for meaning, truth, validity, and manifestation" (ibid.). At this point, it can seem as if MT's solution to how phenomenology and science can cooperate is simply to define phenomenology in this context as a non-transcendental enterprise. If that were the case, it would arguably not be a solution as much as a case of moving the goalposts. What we are after is, after all, a way to naturalize phenomenology that does not simply neglect or erase its philosophical credentials.

MT avoids this objection by pointing to the intimate connection between transcendental and psychological phenomenology, making the latter more of a mediator than a substitute for the former in the envisioned phenomenology-science exchange. Although different from transcendental phenomenology in that it remains within the natural attitude, investigating consciousness as a region of the objective world rather than as a condition of possibility for that world, phenomenological psychology has the potential to *lead to* transcendental phenomenology if pursued in a radical and precise enough manner (ibid., p. 157). In approaching a comprehensive understanding of consciousness as non-reduced phenomenon, phenomenological psychology will eventually be prompted to acknowledge consciousness' transcendental significance. In other words, it seems that the line between psychological and transcendental approaches to consciousness is not so easy to draw after all. On the contrary, on this view, "psychology qua the study of consciousness *contains a transcendental dimension* and is ultimately *part of* transcendental philosophy" (Zahavi, 2017, p. 159; my emphases). This connection between the transcendental and the psychological is also acknowledged by Merleau-Ponty, who states that "the transcendental attitude is already implied in the psychologist's descriptions" (2012, p. 60), even going so far as to label the relationship one of "interpenetration" and "mutual envelopment" (1964, p. 73). The question for MT, however, is how it can subscribe to this way of understanding the relationship between phenomenology and science without sacrificing any of its other commitments. That is, while the transcendental–psychological connection sketched here surely makes more sense of the possibility of mutual enlightenment between transcendental phenomenology and science, it simultaneously hints at a diffusion of the border between the two—a border MT is supposed to leave unquestioned. What does it mean to let "the very conceptions of naturalism and transcendental analysis remain unaffected" (Zahavi, 2017, p. 163) in light of these considerations?

This tension seems to only become more pressing upon further interrogation of this view. Let me draw attention to three points that illustrate this, and which I believe pulls MT closer to IV. First, a possible objection to the view that a productive exchange between phenomenology and science is possible is that there is a mismatch between the *a priori* status of transcendental reflection and the *a posteriori* nature of empirical findings. How can *a priori* insights inform *a posteriori* sciences, or vice versa? Zahavi responds to this by drawing attention to Husserl's view that *a priori* phenomenological insights are not immune to corrections in light of new evidence, but rather "always possess a certain provisionality, a certain presumptiveness" (2017, p. 155). "Our *a priori* knowledge," he elaborates, "is, in short, fallible; if we come across putative empirical counterexamples to our alleged eidetic insights, they need to be taken seriously and cannot simply be dismissed as irrelevant" (ibid.). What we have here, then, is a view of the phenomenological *a priori* not only as *fallible* but as potentially challengeable by *empirical* findings. Now, this idea of revisable transcendental insights is prepared already by what we saw in the previous section regarding the regressive nature of phenomenological transcendentalism. After all, if the task is to start from actual experiences and clarify their constitutive structures, then discoveries that prompt revisions of one's earlier articulations are always a possibility. Still, it is not clear how scientific findings can work as counterexamples to eidetic insights on MT's model.

Second, it is important to note that while the prime example of science in MT's model of phenomenology–science cooperation is phenomenological psychology, which is concerned with a non-reductive understanding of consciousness and takes first-person experience as its point of departure (Zahavi, 2017, pp. 157, 159), this does not mean that "empirical findings" in this context is limited to descriptions of first-personal consciousness as such. Among the empirical sciences that Zahavi mentions as most promising for engaging in productive exchange with phenomenology, we find disciplines such as anthropology, psychopathology, and developmental psychology (2017, p. 152)—all of which, notwithstanding their non-reductive, person-directed nature, at least in part rely on third-personal observations of bodily behavior and its worldly (material and cultural) conditions. Hence, if it is right that findings in these domains "might be taken up by, and consequently influence or constrain, an analysis of transcendental subjectivity" (ibid, pp. 159–160), we need a way to make sense of how third-person perspectives can play this role.

Third, for this transcendental–psychological exchange to work, one must at least admit that the two domains are dealing with subject matters that are closely enough related for them to be relevant to each other. This is staunchly rejected by Gardner, who insists on an "absolute, non-epistemological distinction" between the phenomenal and the objective body (2015, p. 298). This distinction, however, seems to be put into question by Zahavi's Husserl-inspired MT:

> the relation between the transcendental subject and the empirical subject is for Husserl not a relation between two different subjects, but between two different self-apprehensions. The transcendental subject and the empirical subject is but one subject, though viewed from different perspectives. The transcendental subject is the subject in its primary constitutive function. The empirical subject is the same subject, but now apprehended and interpreted as an object in the world (2017, p. 158).

According to MT, then, the separate domains of phenomenology and psychological science are just two different ways of approaching the same subject. While this undoubtedly helps make sense of the relation between the transcendental and the psychological sketched above, it also seems call for a philosophical framework beyond what MT offers. That is, how can a proposal in which "nothing [. . .] entails or necessitates the need for a more fundamental rethinking of the relation between the constituting and the constituted" (ibid, p. 163) make room for the idea that the (constituted) empirical subject is the *same* as the (constituting) transcendental subject?[14] One would think that, without a fundamental rethinking of this relation, the two subjects could *not* be the same, since they would always find themselves at opposite poles of the constitutive correlation. At the very least, the idea of identity between the two subjects would not represent the solution to a problem as much as a problem in itself, as Husserl acknowledged when he in *The Crisis of the European Sciences and Transcendental Phenomenology* described what has come to be known as "the paradox of subjectivity"[15]: "The difference between empirical and transcendental subjectivity remained unavoidable; yet just as unavoidable, but also incomprehensible, was their identity. I myself, as transcendental ego, 'constitute' the world, and at the same time, as soul, I am a human ego in the world" (1970, p. 202). In that work, Husserl's way out of the paradox seems to have been, in Anthony Fernandez' words, "a complete dehumanizing and decontextualizing of the transcendental ego" (2015, p. 294), thus ultimately denying the identity between the empirical and transcendental subject after all. "In the epoché and in the pure focus upon the functioning ego-pole [. . .]," Husserl says, "*nothing human is to be found*, neither soul nor psychic life nor real psychophysical human beings; all this belongs to the 'phenomenon,' to the world as constituted pole" (1970, p. 183; my emphasis). I am not saying that MT necessarily is committed to accept this consequence, but it surely highlights the difficulty of stating the identity of the transcendental and empirical subject within this more conservative transcendentalist position.

Putting together the above considerations, we get a view that says that scientific approaches to consciousness have a transcendental dimension, that *a priori* transcendental analyses are vulnerable to change in light of third-personal empirical evidence, and that the transcendental and empirical subject ultimately is the *same* subject. As we have seen, these features of MT seem to put pressure on its commitment to preserve the classical notions of "transcendental" and "natural." What is missing here seems to be precisely what is offered by our second alternative, the Integrationist View: a model of how the transcendental and empirical aspects of consciousness are *integrated*, so as to make adequate sense of a mutually enlightening relationship between phenomenology proper and science.

As mentioned, one central feature of this view is the aim to rethink the concept of nature in a "phenomenologized" fashion. Of course, "nature" never had a clear and uncontroversial meaning in the first place, so what concept is it more specifically that we are asked to rethink here? Briefly put, [16] it is the *objectivist* concept, which neglects that objects are always accessed by a subject and moreover eliminates anything that is assumed to be mere "products" of human subjectivity (meaning, quality, normativity, etc.) from its picture of the real.[17] While classical transcendentalism tends to be critical of "expansionist" forms of objectivism that purports to shape all forms of thought in its own image, it has generally left objectivist naturalism untouched insofar as it is understood to be a necessary presupposition for the sciences. The assumption that all of natural science is committed to such a position is at least part of the reason for transcendental phenomenologists' long-standing insistence on operating in an autonomous intellectual domain. IV, in contrast, calls for an uprooting of this view of nature altogether, toward one able to incorporate consciousness' transcendental status and the reality of phenomena such as subjectivity, meaning, and normativity. Zahavi points to Thompson's *Mind in Life* as the "currently most comprehensive attempt" at developing such a view (2017, p. 164). In his own words, Thompson's project starts from "a recognition of the transcendental and hence ineliminable status of experience" and aims toward "a different kind of approach to matter, life, and mind from objectivism and reductionism" (2007, p. 87). Central to this approach is the thesis that there is a *deep continuity* pertaining to the organizational structures of mind and life (2007, pp. 128–129).[18]

This leads directly to IV's second defining trait—the transformation of transcendental philosophy from an isolated to a more pluralistic and cooperative enterprise. Where IV's

[14] Compare the quote here to Zahavi's claim in a different text, that "Husserl's phenomenology is characterized by its attempt to modify the static opposition between the transcendental and the mundane, between the constituting and the constituted" (2010, p. 15). It is outside the scope of this paper to explore whether such a modification is compatible with refraining from "a fundamental rethinking," but there at least seems to be a tension at the surface here.

[15] See, e.g., *The Paradox of Subjectivity* (Carr, 1999).

[16] A lot can be—and have been—said about the notions of "nature" and "naturalism" in the context of naturalizing phenomenology. Here, I'm I am limiting myself to a simplified definition, but see, e.g., [Vanzago (2012); Roux (2013), and Reynolds (2018)] for more thorough discussions.

[17] As Zahavi (2013, p. 33) has noted, replacing physicalist/reductionist naturalism with an "emergentist" version alone is not sufficient to ease phenomenology's relation to naturalism, for the latter might still be committed to objectivism.

[18] "Mind is life-like and life is mind-like," as he puts it (ibid., p. 128).

rethinking of nature consists in making room for the constitutive organization of structures of meaning and subjectivity as natural phenomena, its rethinking of the transcendental consists in understanding transcendental reflection as *part of* and directed at nature thus conceived. To make this more concrete, consider how *embodiment* is both a transcendental condition for our openness to the world *and* entails biological existence. (As such, it is a crystallization of the paradox of subjectivity). From a more traditional perspective, the transcendental and the biological would seem to be completely unrelated. IV's conceptual transformations, however, holds the promise of a comprehensive ontology upon which these might be seen as mutually enlightening and constraining perspectives. Zahavi's presentation of this trait of IV seems to point in the same direction. For instance, he cites the suggestion of Roy et al. (1999, p. 61) that "Husserl's and Merleau-Ponty's investigations of the lived body focus on a locus where 'a *transcendental analysis and a natural account are intrinsically joined*'" (2017, p. 164). Furthermore, he draws attention to Merleau-Ponty's call for us to "search for a dimension that is beyond both objectivism and subjectivism" where we would not have to "choose between an external scientific explanation or an internal phenomenological reflection" (ibid, p. 165). This dimension, I will try to show below, is in Merleau-Ponty's first works illuminated through the notion of structures of behavior and—pace Gardner—the preobjective perceptual field. As we will see, this rethinking of transcendental philosophy comes with two significant adjustments relative to its previous form: a step away from first-personal phenomenology, in the sense that the constitution of givenness is no longer an act that is manifest only to the subject of the given, and a recognition of the significance of contingency for its project.

In the remainder of this paper, I will try to show how Merleau-Ponty, despite Gardner's claims to the contrary, offers a promising starting point for developing the IV.[19] An important reason for why this is not noticed in Gardner's reading is that it overlooks two key (and interrelated) factors: (1) the significance of the ontology of *structure* developed in *The Structure of Behavior* (1942/1963) for understanding Merleau-Ponty's overall project[20] and (2) how Merleau-Ponty in the *Phenomenology* limits and transforms transcendental philosophy with his understanding of the phenomenological method. In the next section, I elaborate on the first of these, before I move on to discuss Merleau-Ponty's methodology in *Rethinking Transcendentalism: the Limits of Transcendental Reflection*.

RETHINKING NATURE: STRUCTURES OF BEHAVIOR[21]

What is the "ontology of structure?" Most generally, it is a view of consciousness as an *embodied and expressive* mode of existence that is ontologically prior to the subject–object dichotomy. As Merleau-Ponty puts it in *Structure*'s preface,

> taken in itself, [the notion of behavior] is neutral with respect to the classical distinctions between the 'mental' and the 'physiological' and thus can give us the opportunity of defining them anew. [. . .] By going through behaviorism [. . .] one gains at least in being able to introduce consciousness, not as psychological reality or as cause, but as *structure* (1963, pp. 4–5).

The most important implication of this view for our purposes is that neither the third-person approach of science nor the first-person approach of transcendental phenomenology alone can claim privileged access to the being of consciousness. How does this follow? While I will not be able to give a full account of Merleau-Ponty's ontology of structure here, I will point to a couple of key elements that motivate this conclusion. Let us start by considering a claim Merleau-Ponty makes with regard to what he calls "vital forms," the kind of structure of behavior paradigmatic of non-human animals:[22]

> the reactions of an organism are understandable and predictable only if we conceive of them, not as muscular contractions which unfold in the body, but as acts which are addressed to a certain milieu, present or virtual: the act of taking a bait, of walking toward a goal, of running away from danger (1963, p. 151).

Trivial as it may seem, this observation is of crucial philosophical importance, in the sense that it is a clear illustration of the above-mentioned diffusion of the dichotomy between the subjective and objective. What it says is that, even when approached from a third-person perspective, the behavior of living organisms is expressive of what we might call a "subjective dimension," in the form of displaying a relationship to the world as significant *for* the organism in question. Put differently, Merleau-Ponty is here describing a phenomenon where the "internal" (significances *for* the organism) is expressed in the "external" (observable behavior). The "subjective" or "internal" as understood here is thus not some kind of "extra" feature added upon purely objective movements; it is their structure, or *form*,

[19] "If the promised synthesis of transcendentalism and naturalism could be made plausible independently—no mean feat—then it would furnish the basis for a reconstructive interpretation of Merleau-Ponty, but it is not in Merleau-Ponty's own line of sight" (Gardner, 2015, p. 318).

[20] Admittedly, Gardner does mention the continuity between *Structure* and *Phenomenology* as something that supports the Psychological Interpretation he argues against (2015, p. 296). However, he seems to think that this continuity is only apparent, and limits his attention to *Structure* to a couple of brief remarks describing it as a work of "holist, anti-reductionist thought [. . .] much of which reads like a philosophy of psychology" (ibid.).

[21] Kee (2020) too draws attention to the significance of *Structure* for the project of naturalizing phenomenology. As far as I can see, our approaches are largely aligned, although our emphases are somewhat different. One of Kee's key claims is that a phenomenological reduction is undertaken already in *Structure*'s engagement with psychological considerations, which prompts a shift of attention toward organisms' perceived world. I do not dispute this, but—as we we'll see toward the end of this section—I will argue that a more explicit turn toward (transcendental) phenomenology is motivated by a tension that crystallizes toward the end of *Structure*, and that this is what sets the stage for the *Phenomenology*'s project.

[22] For the sake of simplicity, I'll use vital forms to represent the notion of structures of behavior here. It is, however, important to be aware that Merleau-Ponty locates human behavior at a different level of organization from that of non-human animals. In short, human behavior is not merely oriented with regards to vital/biological needs, but is situated within and directed toward an *intersubjective* world with symbolic significances.

and as such, as I will say more about soon, it is integrated with its "parts" in a relationship of codeterminacy. At its core, the ontology of structure is the view that this embodied-expressive *integration* of subjectivity and objectivity (or the first and third person) is consciousness' primordial mode of existence and hence the ground from which the notions of the mental and the physiological are abstracted.[23]

Now, admittedly, the claim in the above quote is not so much ontological as it is epistemological, i.e., it is telling us how organisms' behavior must be *conceived* in order to be understandable, rather than establishing that organisms ultimately *are* one way rather than another. Here, then, we are confronted with a challenge to my claim above that the subjectivity exhibited by organisms' behavior is not an additional, separable feature—for, can it not be the case that the "sense" we see in living behavior is merely a result of our projections as observers?[24] This challenge is a decisive moment in the dialectic toward the Integrationist View, motivating, as it does, a return to the position of the philosopher or scientist *qua* the subject seeking to understand consciousness' place in nature: Can it not be the case that the sense displayed in the behavior of living organisms is merely the result of *our mode of understanding or perceiving*, and not something that is "really there," in nature? This is the cue for the transcendental philosopher to step onto the stage: the focus have now shifted from the nature of mind and life to how the phenomena of structures of behavior—in this context, vital forms—are constituted as phenomena *for* consciousness. The perspective on consciousness we have entertained so far in this section has been *transcendentally naive*—it has been that of an "external spectator" leaving its own status *as* spectator unquestioned. As many, including Merleau-Ponty himself, have noted, this is the perspective from which most of *Structure* is written[25]; only in the last chapter of that book do we see a shift begin to take place toward a "transcendental" perspective. Let us, however, leave the execution of this shift on hold for a moment, while we let the naive spectator provide us with some more flesh on the bone of the ontology of structure.

The ontology of structure consists in taking the organism as a whole, in its dynamic interactions with its environment, as an irreducible "unit" of nature. Irreducible, because the existence and function of any smaller "part" of this unit (such as physiological features) depends on it being a part of this greater whole, just as the whole in turn depends for *its* existence on the existence and functioning of its parts. As such, structures of behavior are characterized by what Evan Thompson labels "dynamic co-emergence," meaning "that a whole not only arises from its parts, but the parts also arise from the whole. Part and whole co-emerge and mutually specify each other" (2007, p. 38). While this sort of part–whole relationship can be found also in some non-living physical structures, the structure of living beings is further characterized by having an equilibrium that depends upon "virtual" conditions—that is, conditions produced by the organism itself, and which hence do not exist independently of it (Merleau-Ponty, 1963, p. 145). In Thompson's terminology, living organisms are *autonomous* systems (2007, p. 37)— systems that themselves generate and maintain the processes necessary for continuing their existence as such. As a concrete case in point, consider the way in which your existence, as a bodily being, is generated by metabolic processes that in turn are maintained only insofar as you interact with your surroundings in a certain way—seeking food when you're hungry, safety when you're scared, and so on. Neither these vital significances of your external world ("food," "safety") nor the metabolic processes of your cells are things that can exist independently of you as a holistic structure of behavior—they are brought forth *by* this structure while simultaneously being among the conditions necessary for the maintenance of the same structure.

There are two main reasons for why these points regarding the autonomous and dynamic coemergent nature of structures of behavior are important for our purposes. First, they provide us with a helpful framework for making sense of what I earlier called the "subjective dimension" of living structures. Second, they enable us to see how consciousness, *qua* dynamically coemergent structure, is *vulnerable and contingent*—a point that is key to understanding how the Integrationist View sees the relationship between phenomenology and empirical science.

Starting with the first of these, consider how the notion of autonomous systems accounts for the existence of the three interrelated phenomena of (1) selfhood or individuality, (2) a world or environment with a certain relevance or sense *for* the system, and (3) normativity concerning the system's state and interactions.[26] Through generating and maintaining itself, the system produces itself *as* self or individual by distinguishing itself from its surroundings. By way of the same process, the surroundings gain a sense or relevance for the system in light of its project of self-generation and self-maintenance. Given that a certain functioning both of the system's internal organization and of its interactions with its surroundings are of literally existential significance, the emergence of autonomous systems is simultaneously the emergence of a form of normativity pertaining to the system in question; certain states and interactions are more *preferable* than others for the organism

[23]To avoid any misunderstanding, let me emphasize that the acknowledgment of an "internal" dimension of behavior here means that we, despite the focus on behavior, are going beyond traditional (objectivist) behaviorism. Merleau-Ponty's point in the above quote is that the objectivist stimulus—response approach of behaviorism—which neglects, as Kee puts it, "the perceived world of the animal itself," thus failing to recognize "behavior and situation as *internally* related structures with a unique logic" (2020, p. 19)—is unable to adequately understand behavior.

[24]Let me note, without being able to argue extensively for it here, that a concession to such a separation between epistemology and ontology is problematic, in the sense that it ultimately leads to an unsustainable skepticism. For someone who challenges the epistemology—ontology separation, see, e.g., Bhaskar (1978, pp. 36–45) or Taylor (1997). Although it is not explicitly addressed in the main text, the claim that there is a sort of meaning *for* living organisms as well as immanent *in* their behavior, is in effect a subscription to teleological view of life. Thus, the epistemological challenge here can be more specifically framed as one of teleonomy (living organisms behave/must be understood *as if* purposive) vs. teleology (living organisms *are* immanently purposive). See Weber and Francisco (2002) for an informative overview of this issue as well as an argument in favor of the latter position.

[25]For instance, in one of *Phenomenology*'s footnotes, he distinguishes *Structure* as concerned with consciousness "seen from the outside" from *Phenomenology* as concerned with consciousness "seen from within" (2012, n18 p. 535).

[26]Roughly the same points can be found in Thompson (2007, pp. 73–74).

in light of its project to keep on existing. In short, a living organism's structure of behavior is expressive of a network of relations (of dependence, interests, understanding, etc.) between the organism as individual and its environment, brought forth by and meaningful in light of the self-concern of the organism as a whole. This is what accounts for the "subjective dimension," or form, of living organisms' behavior.

As hinted above, a crucial consequence of this view is that it amounts to what we might call a "deprivatization" of consciousness. That is, as understood here, consciousness resides in embodied-expressive behavior and is as such not exhausted by its first-person access to itself but is publicly available. This view is expressed by Thompson when he says that

> The intentional arc and being-in-the-world overall are neither purely first-personal (subjective) nor purely third-personal (objective), neither mental nor physical. They are existential structures prior to and more fundamental than these abstractions (2007, p. 248).

"The intentional arc" denotes the network of relations between the living organism (or subject) and its world or, in other words, the ways in which the former is situated in and directed toward the latter. To conceive of the intentional arc as an existential structure is to give up on the idea of consciousness as an essentially "inner" mode of being. "The mental," as Merleau-Ponty puts it in *Structure*, "is reducible to the structure of behavior" and "[s]ince this structure is visible from the outside and for the spectator at the same time as from within and for the actor, another person is in principle accessible to me as I am to myself [. . .]" (1963, pp. 221–222). This view is carried on even after the shift from *Structure*'s "spectator perspective" to *Phenomenology*'s "internal" study of consciousness, where Merleau-Ponty already in the preface echoes the citation from *Structure* in stating that

> I must be my exterior, and the other's body must be the other person himself. [. . .] my existence must never reduce itself to the consciousness that I have of existing; it must in fact encompass the consciousness that *one* might have of it, and so also encompass my embodiment in a nature and at least the possibility of an historical situation (2012, lxxvi).

In other words, my existence as consciousness is not limited to my first-person perspective but extends beyond it. Importantly, this should not be understood as an elimination of first-personal experience or a rejection of the idea that each individual enjoys a special sort of "access" to his or her lived experience that is unavailable *as such* to others. The point is that my reality as subject exceeds what I can grasp through my own perspective upon myself, which means that I am not the sole authority when it comes to understanding my own existence. This deprivatization of consciousness suggests a rethinking of transcendental philosophy, which will be further explored in the next section: If the transcendental is the structural organization in virtue of which stuff appears to consciousness in the first place, and consciousness is an existential structure not exhausted by the first person but visible from "the outside," then it seems plausible that transcendental reflection must incorporate this "external" view upon consciousness in order to be adequate.

The second reason mentioned above, concerning the contingency and vulnerability of structures, points in the same direction. If consciousness is an existential and bodily structure characterized by a relation of dynamic coemergence between parts and whole, there seems to be little room for the traditional transcendental trait of *a priori necessity* in its organization. On the contrary, being dependent on the proper functioning of its parts for the maintenance of its mode of existence, the structure of consciousness seems to be susceptible to significant reorganizations in reaction to empirical events. Consider, for instance, the case of pathology. It is tempting to understand pathological subjects on the model of a "normal" human way of being, thematizing the illness as a lack or distortion of individual features or capacities relative to this standard. As we will see in the last section of this paper, however, pathology is better understood if we acknowledge that illness, as Merleau-Ponty puts it, "is a complete form of existence" (2012, p. 110). As such, pathology is not so much a case of absent or disturbed particularities relative to an otherwise intact "normal" structure, as it is a different way of existing altogether, a novel, albeit disintegrated, way of being situated in and directed toward the world.[27]

The takeaway for now is this: Considered as existential structure, consciousness is both "deprivatized" and fundamentally contingent. Both of these seem to suggest that there is an important role for empirical science to play together with phenomenology in illuminating the structures of consciousness. After all, empirical perspectives are required in order to adequately grasp the contingencies of our embodied existence, describing, for instance, how humans' way of relating to the world is affected by bodily injury and traumatic experiences, or what role empirical matters play in childhood development. The question, however, is whether the above reflections have any bearing on the transcendentalist challenge to naturalized phenomenology.

While I think that the ontology of structure in the end will prove to offer what MT lacks (i.e., a framework for understanding how the transcendental subject and the empirical subject can be the *same* subject approached from different perspectives), I am not under the illusion of having convinced the transcendental philosopher yet. After all, recall that the conception of consciousness as structure of behavior that has hitherto been developed is based on the transcendentally naive perspective of an "outside spectator" taking for granted the way in which his or her access to phenomena is constituted or achieved in the first place. Even though I have been advocating a conception of consciousness as an embodied-expressive structure integrating subjectivity and objectivity, thus

[27] In the words of Georges Canguillhem, "disease is not a variation on the dimension of health; it is a new dimension of life" (1978, p. 108). This does not mean that there are no meaningful distinctions to be made between the normal and the pathological, in the sense that the latter cannot be seen as the "worst" of the two. The point is that what is disturbed in the pathological case is ultimately the global behavioral space of the patient, not individual psychological or physiological traits.

challenging standard "objectivist" views of nature, this model is still that of consciousness as an *object in the world*, as a phenomenon *for me* as perceiver and thinker. What the transcendental philosopher is concerned with, remember, is not consciousness as object but rather as our *access* to objects as such. This is probably the explanation for why Gardner so quickly dismisses the significance of Merleau-Ponty's first work in the context of determining his stance toward transcendental philosophy.[28]

There are, however, at least two reasons why this dismissal is too quick. First, it seems possible to argue that the above-mentioned elements of the ontology of structure are in fact relevant to the transcendental project. After all, what I have tried to outline here is a rethinking of nature[29] that acknowledges a dimension of being prior to the subject–object dichotomy, where consciousness is reconceived as essentially manifest in the grammar of behavior, and our bodily being is seen as organized toward an environment of meaning. Thus, although the ontology of structure might have its origin in a third-personal perspective, its ultimate consequence is a diffusion of the distinction between first- and third-personal perspectives—organisms' perceived, meaningful environment is exhibited in the behavior by which it is enacted, and must be incorporated as such by the ("third-personal") scientist in order to be adequately understood. Since observable bodily behavior in this way expresses the constitution of an environment of meaning, it seems possible that it also bears clues of transcendental significance, disclosable by scientific perspectives. In the next section, we will see how this transcendental significance of the notion of structure is motivated also by the "internal" perspective of the *Phenomenology*.

Second, and on a more scholarly note, Gardner seemingly ignores how Merleau-Ponty in the last pages of *Structure* sets the stage for the *Phenomenology's* transcendental project. That is, there is a tension running through *Structure* that is brought to the fore in its last chapter and which seems to be the motivation for at least some core parts of the *Phenomenology*. This tension is related to the challenge mentioned earlier, regarding the relation between the contributions of our mode of understanding and the embodied-expressive *sense* characteristic of the holistic structure of living organisms' behavior. The problem, as Toadvine notes, is that the ontology of structure is based upon how the behavior of living beings *appears* as meaningful wholes to a subject, thus giving "the impression that [it] involves a return to idealism, since every structure would have consciousness as its essential correlate" (2009, p. 38). We are thus confronted with the possibility that the sense we disclose in nature belongs only to nature *for us*, that it is a product of our human significations.

Merleau-Ponty's diagnosis of this problem points directly toward his project in the *Phenomenology*. What leads us toward the idealistic conclusion is that we identify structures with *significations* dependent upon our human conceptualizing capacities, thus privileging, as Toadvine puts it, "the perspective of intellectual consciousness" (ibid.) as our access to the world. The problem can be avoided, however, if we acknowledge that intellectual consciousness is derivative from *perceptual* consciousness, and hence "return to perception as a type of original experience in which the real world is constituted in its specificity" (Merleau-Ponty, 1963, p. 220). This call for a *phenomenology of perception* is made in explicit opposition to traditional, Kantian transcendental philosophy, which sets the contribution of the categories of the understanding center stage. Given that this intellectualist theory is not acceptable, Merleau-Ponty famously states in *Structure*'s very last page, "it would be necessary to define transcendental philosophy anew in such a way as to integrate with it the very phenomenon of the real" (1963, p. 224). Here, we have a clear formulation of the path forward for Merleau-Ponty's thought, i.e., to *redefine* transcendental philosophy in a way that does justice to the reality of structures. While the shift from intellectual to perceptual consciousness, which Gardner too acknowledges, is a crucial part of this redefinition, an equally important factor is the way in which this entails a methodological integration of phenomenology with scientific perspectives. How are we to understand this methodology?

RETHINKING TRANSCENDENTALISM: THE LIMITS OF TRANSCENDENTAL REFLECTION

The *Phenomenology* immediately picks up the thread from *Structure*, addressing the nature of the methodology that is to be employed in its "return to perception." Thus Merleau-Ponty starts, in the very first sentence, with the question "What is phenomenology?" (2012, p. lxx). According to Gardner, Merleau-Ponty in his response "*avows a commitment to* phenomenology conceived as 'a study of essences,' 'a transcendental philosophy,' 'a rejection of science'" (2015, p. 304; my emphasis), thus confirming his own transcendentalist reading. Looking at what Merleau-Ponty is actually saying in the relevant passage, however, this is a far too strong claim, highlighting only one side of what is really presented as *tensions* found within phenomenology. That is, the claim that phenomenology "is the study of essences" is immediately followed by the qualification that "yet [it] also places essences back within existence and thinks that the only way to understand man and the world is by beginning from their 'facticity'" (2012, p. lxx). Furthermore, while phenomenology is "a transcendental philosophy [. . .] *it is also* a philosophy for which the world is always 'already there' prior to reflection" (ibid.; my emphasis). Lastly, after stating that phenomenology attempts to describe experience "such as it is, without any consideration of its psychological genesis or of the causal explanations that the scientist [. . .] might offer of that experience," Merleau-Ponty points out that in the last works of Husserl one also finds the notion of a "genetic phenomenology" (ibid.). Ending the paragraph

[28] See footnote 24.

[29] Admittedly not nature as a whole, but more specifically in the form of living organisms. However, as Morris (2018) has recently tried to show, this approach to the structure of life might ultimately help us toward a more fundamental rethinking of nature.

by referring to these as "contradictions,"[30] it is clear that he is here not *avowing a commitment* to a specific conception of phenomenology as much as he is, as Reynolds puts it, acknowledging "a constitutive methodological disunity at the heart of phenomenology" (2017, p. 87). Rather than providing an answer to the initial question, then, Merleau-Ponty is here offering a further elaboration of the *difficulty* of providing such an answer.

Given the significance of this "methodological disunity" for understanding the nature of Merleau-Ponty's project, it is surprising that the question of the phenomenological method is not addressed at all in Gardner's paper.[31] If Merleau-Ponty is a transcendental philosopher, then by what means does he access the transcendental domain? Here, his claim, which we briefly touched upon in *The Transcendentalist Challenge and Varieties of Transcendentalism*, that "the most important lesson of the reduction is the impossibility of a complete reduction" (2012, p. lxxvii) becomes relevant. What does it mean?

The way to make sense of this at first glance enigmatic statement is to see it as signaling an immanent critique of transcendental philosophy, in the sense that it represents a case of turning transcendental reflection against itself. That is, starting from the position of the transcendental philosopher aiming to fully grasp the constitutive structures of the world's presence for consciousness, we discover a resistance to our endeavor that ultimately turns out to be an unsurpassable limitation for our project, namely, the fact that we are situated and inextricably involved in a world in ways that can never be exhaustively conceptualized. This, in short, is the fact of embodiment, the concrete, perspectival nature of our existence that makes presence always come at the expense of a certain absence, most simply exemplified by how the visual presence of objects is characterized by the absence of the sides not facing us. The general point here is that, due to our situated, bodily nature, any act of bringing something into view, of achieving presence, or of thematization, is enabled by a background that is "out of view," absent, or unthematized. From our position as transcendental philosophers, this is obviously a problem: It entails that our reflection, which aims to illuminate the enabling conditions of experience, itself depends upon conditions that it *cannot* fully thematize.

In other words, what we learn from the reduction is, negatively, that the presence of the world *resists* our attempt to reduce it to something that can be exhaustively thematized in acts of reflection and, positively, that consciousness is characterized by a primordial and inescapable bond to the world, which is presupposed by all of our more intellectual mental activities. Thus, the impossibility of a complete reduction has implications for our understanding of both the methodological status of phenomenology as well as the ontological status of consciousness.

First, note that the assumption that a complete reduction is *possible* itself rests upon an unquestioned, naive presupposition, namely, that of a subject enjoying full reflective access to the structures constitutive of its openness to the world. Furthermore, the thought that the execution of the *epoché* can provide this sort of access seems committed to the belief that the meaningful presence of the world is reducible to a meaning for consciousness *qua* reflecting subject. Thus conceived, the phenomenological reduction would, as Merleau-Ponty remarks, "be idealist, in the sense of a transcendental idealism that [. . .] strips the world of its opacity and its transcendence" (2012, p. lxxv). This description, I think, fits the sort of idealism—which mistakes perceived form for intellectual significance—that was at the root of *Structure*'s tension concerning the ontology of structure. What we are seeing here, then, is an internal critique of that view: Taking the possibility of its own project for granted, transcendental reflection's search for the presuppositions of experience is blind to *its own* presuppositions. In other words, it neglects that we are not constantly reflecting subjects, but that reflection has a *beginning*, and as such is "a genuine creation, a change in the structure of consciousness [. . .]" (ibid., p. lxxiii). The task thus becomes one of uncovering the origin of reflection and the unreflective ground from which it arises.

This is the task of what Merleau-Ponty calls "radical reflection" (ibid., p. lxxviii), which, in Toadvine's words, is a reflection that "aims to take into account its own immemorial past, its pre-reflective life in nature, as the fundamental condition for its operation as reflection" (2009, p. 53). This, then, is the method of Merleau-Ponty's novel form of transcendental philosophy, distinguished by its aim to uncover the genesis of reflection rather than taking it for granted. How can this be done? I think Morris is on to something when he observes that "who we are as reflectors [. . .] is a much more contingent and empirical question than the naïve [intellectualist] view would allow" (2018, p. 85). In order to see how that is so, consider the ontological implication of the discovery of the impossibility of a complete reduction. Leading us to recognize our inextricable entanglement with the world as embodied beings, the assessment of a complete reduction as impossible is a first step toward establishing "from the inside" what *Structure* did "from the outside;" namely, that the being of consciousness is primordially that of an embodied structure of engagement with and situatedness within a world or environment, not fully graspable from this structure's "subjective" point of view. This is another essential turning point in the dialectic toward the Integrationist View: The objection that the notion of consciousness as structure of behavior is transcendentally naive is here countered with the observation that the reflecting activities of transcendental philosophy *themselves presuppose the philosopher's existence as structure*. In short, just like an adequate third-person understanding of living organisms presupposes recognizing the "subjectivity" displayed in their behavior, an adequate first-person understanding of subjectivity presupposes recognizing it as integrated in a living

[30] "Might one hope to remove these contradictions by distinguishing between the phenomenologies of Husserl and Heidegger?" (2012, p. lxx). See Zahavi (2008) for a closer discussion of these remarks from the *Phenomenology*'s preface.

[31] As far as I can see, Gardner (2015) mentions the phenomenological method only twice, first in the context of presenting the Psychological Interpretation (p. 296), and later (p. 304) when interpreting a claim Merleau-Ponty makes in *The Visible and the Invisible* (1968). He does, however, never discuss or clarify what he takes this method to be.

body's "deprivatized" and contingent mode of existence as structure.[32]

Let us revisit the phenomenon of contingency now that we have established the connection between transcendental philosophy and structures of behavior. Transcendental philosophy, remember, usually aims to identify the necessary, constitutive structures of experience. Given what we saw in the previous section, regarding the fundamental contingency of consciousness understood as structure, can we still talk about any sort of necessity pertaining to its organization? We might, in this sense: For any mode of existence, there will be processes, structures and features that are necessary for it to maintain its specific way of being situated in and directed toward the world *in the way that it currently is*. That is, as a dynamically coemergent structure, every aspect of our embodied being is in some sense necessary for our holistic form of existence to remain as it is. Being dependent upon empirical contingencies and as such vulnerable to change, however, such "necessary" features of consciousness understood as structure are necessary only in a limited, relative sense. This, I take it, is Merleau-Ponty's point when he states that

> It is impossible to distinguish in the total being of man a bodily organization that one could treat as a contingent fact and other predicates that necessarily belong to him. Everything is necessary in man, and, for example, it is not through a simple coincidence that the reasonable being is also the one who stands upright or who has opposing thumbs—the same manner of existing is expressed in both of these cases. And everything is also contingent in man in the sense that this human way of existing is not guaranteed to each human child through some essence acquired at birth [. . .] (2012, p. 174).

This unorthodox conception of necessity—which refers to the constitutive, yet contingent form of embodied human existence—seems to include any detail that contributes to our total way of being in the world as such. Ultimately, if put "back into my living body" (i.e., if seen as parts of my holistic embodied-expressive existence) even "my ears, my nails, and my lungs [. . .] will no longer appear as contingent details" because "[t]hey are not indifferent to the idea of me that others form, they contribute to my physiognomy or to my style" (2012, p. 455). "Physiognomy" and "style" here refer to my existence as embodied-expressive structure. Notice that this appeal to how I am present to *others* is in line with the above-mentioned deprivatization of consciousness—my existence as subject is not exhausted by my first-personal access to myself but comprises my existence as appearance in the world, available to other perspectives.

It is difficult to see how this notion of necessity we have discovered here is compatible with Gardner's claim that "the conditions that [Merleau-Ponty's] phenomenology uncovers are intended to be in the true and genuine sense transcendental, i.e., *a priori* and necessary, and non-identical with empirical, contingent, or mundane states of affairs" (2015, p. 300). On the contrary, what we seem to have now is an *integration* of necessity with contingency, in the sense that the "necessity" pertaining to consciousness as structure is merely its holistic organization of contingent details on which it in turn depends.[33]

If the transcendental has now become the holistic organization of a deprivatized and contingent consciousness *qua* existential structure, then the idea of transcendental philosophy as completely indifferent to scientific matters seems hard to defend. That is, if we want to understand how reflection can arise from our existence as structure, we have to involve perspectives that can illuminate structures of our being not accessible from our point of view as self-reflecting philosophers.

Let us return to the position of the first-person phenomenologist in order to get a better grip on the task at hand. While the phenomenological reduction, as we have seen, cannot consist in a full bracketing or suspension of our attitude toward the world around us, it is nonetheless a productive undertaking in the sense that it loosens "the intentional threads that connects us to the world in order to make them appear" (2012, p. lxxvii). In other words, the reduction enables us to appreciate the complexity of our dependence on and directedness toward the world—the complexity, that is, of the intentional arc. What is thus revealed is a field of preobjective, indeterminate, or *ambiguous* phenomena—the phenomenal, or *transcendental*, field. This field is transcendental in the sense that it is the always presupposed ground for our thoughts and reflections. As we saw in *The Transcendentalist Challenge and Varieties of Transcendentalism*, ascribing this role to a *preobjective field* represents a significant shift from traditional transcendental philosophy. Furthermore, the preobjective nature of the transcendental field means that, while being an enabling factor *for* reflection, it is never fully graspable by the objectifying acts *of* reflection. The word "field," says Merleau-Ponty, "signifies that reflection never has the entire world [. . .] spread out and objectified before its gaze, that it only ever has partial view and a limited power" (2012, p. 62). The reflecting subject inevitably finds herself always already situated *within* the phenomenal field, presupposing this unreflective bond in all acts of reflection, in the sense that thought only ever gets started against the background of something *unthought*. The phenomenal field, then, is fundamentally *ambiguous* or *indeterminate*, since it resists reflection's demands for determinacy and clarity by always escaping its full grasp, lending itself to an indefinite number of alternative—perhaps conflicting—acts of determination.

The task of radical reflection is to illuminate this field and the intentional arc that sustains it, with the aim to understand how reflection emerges in the first place. What is clear from what we have seen above is that a philosopher's reflections alone are not up for this job. Given the complexity of the intentional arc, which "ensures that we are situated within [. . .] our past, our future, our

[32] A key point here is how our self-understanding originates in and depends upon intersubjectivity (intercorporeality), and further how aspects of the self unavailable through self-reflection can be made "visible" through the other's mimetic responses to my behavior. I will unfortunately not be able to further elaborate this here.

[33] Andrew Inkpin draws attention to similar points regarding necessity in the *Phenomenology* (2017, p. 40). He, however, uses this (among other considerations) to conclude that Merleau-Ponty's project does not warrant the name "transcendental." While I agree that this is the right conclusion given Inkpin's criteria for what counts as transcendental philosophy, I think the historical fluidity of this notion makes it possible to rather see Merleau-Ponty as transforming the idea of the transcendental.

human milieu, our physical situation, our ideological situation, and our moral situation" (Merleau-Ponty, 2012, p. 137), this task is multifaceted and demands a variety of approaches. This, I think, is how we should understand the significance of Merleau-Ponty's extensive engagements with empirical science throughout his *oeuvre*—that is, as integrating these non-philosophical perspectives into his project of radical reflection. Note that this is a direct contradiction of Gardner, who reads Merleau-Ponty as ultimately reaching a purely transcendental domain, strictly independent of scientific considerations. What we have seen above, on the contrary, is that the idea of transcendental reflection as autonomous and independent from anything other than itself is doomed to neglect its own emergence as reflection and hence fail to adequately perform its task, left rather to spin in a frictionless void[34] of its own creation. As *radical*, transcendental reflection acknowledges its own limitations and seeks to incorporate a plurality of perspectives in its project of uncovering the ground for its own genesis in contingent and deprivatized existential structures of behavior. Thus, rather than being opposed to all efforts of "naturalization," a genuine transcendental phenomenology rather *requires* an integration of phenomenological and scientific perspectives.[35]

"Integration" is here not meant to entail complete alignment of the transcendental with the scientific in all respects—after all, their aims are often distinct, and communication between them, while desirable, is never guaranteed. What it means is that they are in a relationship of "interpenetration," as Merleau-Ponty would say: Each have the potential to gain something from the other and should be pursuing this possibility given the fact of their common origin and their participation in the same pluralistic field of nature. Admittedly, this claim remains empty as long as we have not seen such an integration actualized on a concrete case. This is perhaps the most important implication of the Integrationist View: It is actualized *in* integration, in the sense that it is first in engagement with the concrete and particular that its content is adequately articulated. Thus, let us finally turn to see how Merleau-Ponty's engagement with the Schneider case is an example of Integrationist naturalized phenomenology in practice.

INTEGRATION IN ACTION: THE SCHNEIDER CASE

The scientific research that figures most extensively in the *Phenomenology* is Adhémer Gelb's and Kurt Goldstein's studies of neurological pathology in World War I veterans, in particular their observations of the patient called "Schneider," who had been struck by shrapnel from a mine to the back of his head, causing severe injury to his brain.[36] Although Merleau-Ponty engages with various aspects of the Schneider case throughout several chapters of the *Phenomenology*, I will here limit my discussion to Schneider's pathology as it is presented in Part 1, Chapter III, "The Spatiality of One's Body and Motricity." There, the main concern is the relation between "abstract" and "concrete" movements. A distinction between these forms of movements is suggested by the fact that Schneider is unable to point to areas of his body when asked to do so, although he is perfectly able to grasp or touch the same areas if those movements are called for by the immediate, concrete situation, for instance when bitten by a mosquito (2012, p. 106). This distinction is applicable also to another curious abnormality displayed by Schneider: Although he performs the tasks of his work without difficulty when in the actual situation of his working place, having the required instruments at hand, he is unable to *imitate* the same movements without elaborate preparations, having to, so to speak, actively "place" his whole body virtually within the concrete situation of his working place (ibid, p. 112). Similarly, he is unable to move his hand into a military salute without assuming a whole military-like posture, producing the concrete situation where such a movement is called for (ibid.).

What we have here is a collection of empirical descriptions of patterns of pathological behavior that emerged after Schneider's accident. The question now is, what is the significance of these facts for phenomenological philosophy? An appealing yet ultimately too simple way to understand Merleau-Ponty's dealings with the descriptions of Schneider's pathology is that he brings a preestablished philosophical framework to bear on a concrete case, corroborating his ontology of structure by showing its supremacy over alternative ways of accounting for the facts. This way of looking at it can be motivated by reconstructing the trajectory of Merleau-Ponty's thought in the relevant chapter as follows: Revealing the shortcomings of intellectualist and empiricist explanatory strategies, which try to reduce Schneider's pathology to a malfunctioning pertaining to either the causal processes of the physiological body or to the representational capacities of the mind, Merleau-Ponty shows how Schneider's disorder can best be understood through the model of existential structures. The problem of whether to account for the disorder as *either* physiological or psychological is thus overcome through the idea of the living body as a structure of behavior: "The motor disorders in cerebellar injury cases and those of psychic blindness can only be coordinated if the background of movement and vision is defined not by a stock of sensible qualities, but by a certain manner of articulating or of structuring the surroundings" (2012, p. 117).

In saying that this way of rendering Merleau-Ponty's use of the Schneider case is misleading, I do not mean that it is completely false. After all, it is clear that Merleau-Ponty believes his non-reductive, phenomenological approach contributes to a better understanding of Schneider's pathology. As he says,

> Behavior can only be grasped by [. . .] the type of thought that takes its object in its nascent state, such as it appears to him who lives it, with the atmosphere of sense by which it is enveloped, and that seeks to slip itself into this atmosphere in order to discover, behind dispersed facts and symptoms, the total being of the subject (2012, p. 122).

[34] To borrow McDowell's (1994) expression.

[35] Reynolds makes the same point: "[Merleau-Ponty's] particular conception of transcendental philosophy [. . .] not only is compatible with a serious and sustained engagement with empirical science but even requires it" (2017, p. 87).

[36] Schneider's illness was initially diagnosed as a case of visual agnosia (Goldstein and Gelb, p. 137), and has later been more narrowly classified as a case of "visual form agnosia" (Farah, 2004, p. 13). I'm I am indebted to Rasmus Thybo Jensen's paper on Merleau-Ponty's engagement with the Schneider case for these references (Jensen, 2009).

What is misleading is the idea that the relation of enlightenment between philosophy and science here only goes one way, from the former to the latter. In order to see how it rather is a case of *mutual* enlightenment, remember first that the task of radical reflection as staked out in the previous section is to illuminate the structures of its own prereflective ground. Now, how can the study of Schneider's illness contribute to this project?

Consider the significance of this empirical case as a concrete example of a radical modification of the intentional arc. It is, in other words, an existence proof of the contingency of human consciousness' situatedness in and directedness toward the world and as such reveals the reality of the integration of the transcendental and empirical.[37] Schneider's brain injury initiated a process of a global restructuration of his mode of existence—his motor and cognitive capacities became organized in a new way, took on a new sense, in order to cope with the challenges arising from the damage. Since his pathology concerns his subjectivity as much as his motricity, he no longer has access to the same *phenomenal field* as normal healthy subjects. In other words, the "necessity" of Schneider's normal human way of disclosing the world was necessary only up until the point a shrapnel hit his head.

Now, as mentioned in *Rethinking Nature: Structures of Behavior* above, this has some important consequences for what we can learn from comparing Schneider with "normal" subjects. As Merleau-Ponty emphasizes, "It cannot be a question of simply transferring to the normal person what is missing in the patient and what he is trying to recover. Illness, like childhood [. . .], is a complete form of existence [. . .]" (2012, p. 110). Just like we cannot understand children's way of existing as an adult form with certain lacks, the relation between the normal and the pathological subject cannot be understood in terms of subtraction or addition of particular functions pertaining to an otherwise identical structure. Thus, we should not take Schneider's pathological behavior as exhibiting either the lack or the presence of particular functions that we can then infer are also present in normal subjects. If Schneider's disorder is not a case of a malfunctioning of any one particular function, but rather the manifestation of a pathological *mode of being*, then the difference between him and normal subjects must be a difference in existential structure—in their global way of engaging with the world.

On this basis, we can use the contrast between Schneider and normal subjects in order to disclose a "transcendental" structure of organization characteristic of the latter's mode of existence. That is, in considering Schneider's inability to easily and immediately move his body accordingly in response to instructions even though he understands them intellectually, we are able to catch a glimpse of a structural moment of our prereflective embodied existence that would hardly have been available through the transcendental philosopher's reflections. Since Schneider, Merleau-Ponty observes, "is missing neither motricity nor thought," but nonetheless displays this inability to perceive "motor significations," "we must acknowledge, between movement as a third person process and thought as a representation of movement, an anticipation or a grasp of the result assured by the body itself as a motor power [. . .] or a 'motor intentionality"' (2012, p. 113). With this notion of motor intentionality, which is neither a purely mechanical physiological process nor an explicit first-personal thought, we have thus discovered a way to conceptualize our being as embodied structures in a way that was not available to us prior to the empirical case of Schneider's disorder.

While this might seem as a case of inferring the presence of a feature in normal cases from a lack in the pathological case, the point is on the contrary that motor intentionality has to do with the *total* organization of the normal human structure of behavior.[38] This is Merleau-Ponty's point when he argues that "'visual representations,' 'abstract movement,' and 'virtual touching' are only different names for a single central phenomenon" (2012, p. 120), or again, that "visual representations, tactile givens, and motricity are three phenomena cut out of the unity of behavior" (2012, p. 121). One way to characterize the structure of behavior of normal subjects as opposed to that of Schneider is to say, as Merleau-Ponty does, that "the normal person *reckons with* the possible, which thus acquires a sort of actuality without leaving behind its place as possibility" (2012, p. 112). In short, we inhabit a world that, in an important sense, is more *open* than that of Schneider, who, we can say, is "trapped" in an environment that does not offer him the same behavioral possibilities as we have. Thus, motor intentionality—the capacity for immediate bodily grasp of significances—is in the normal case a power that characterizes our total mode of being in the world. In other words, it sustains the *intentional arc*—which, to repeat, is what ensures our situatedness within a complex network of natural and symbolic relations, and further "creates the unity of the senses, the unity of the senses with intelligence, and the unity of sensitivity and motricity"—and it is *this*, Merleau-Ponty claims, which ultimately " 'goes limp' in [Schneider's] disorder" (2012, p. 137).

To sum up, if Merleau-Ponty's analysis is correct, radical reflection has here made progress in uncovering some of the conditions that enable it: Motor intentionality has been established as a transcendental power integrated with contingent embodied life and discovered through engagement with a

[37] Fernandez makes a similar point in arguing that Schneider motivates the idea of a "contaminated transcendental": "The *a priori*, ontological structures of the world are contingent precisely because they are contaminated. And the contaminant is the world itself" (2015, p. 296).

[38] Jensen (2009) points to an interesting and significant ambiguity on Merleau-Ponty's part regarding his understanding of the power of motor intentionality. That is, Schneider here seems to be used "in two mutually exclusive ways: motor intentionality is to be revealed both by its perspicuous preservation and by its contrastive impairment in one and the same case" (Jensen, 2009, p. 372). Does Schneider exhibit an intact normal form of motor intentionality, or is rather his illness a case of a distortion of this same function? Given the holistic approach to structures of behavior that I 've have advocated above, the idea that *the very same* power of motor intentionality is at work both in normal subjects and in Schneider must be rejected. Rather, if we are to talk about "motor intentionality" in both cases, it must be two different *forms* of motor intentionality—one sustaining a "normal" human organization of existence and one facilitating the maintenance of a more disintegrated, pathological mode of being. While the textual evidence for a contradiction on Merleau-Ponty's part is surely real, Jensen's conclusion that "[t]he best way to avoid the contradiction is to accentuate the differences between the concrete actions of the patient and the corresponding actions performed by the normal person [. . .]" (2009, p. 387) supports this reading.

scientific account of an empirical case. As such, Merleau-Ponty's engagement with the Schneider case has yielded a result relevant to both science and transcendental phenomenology and is thus a clear case of the Integrationist View in action.

CONCLUSION

My aim in this paper has been to propose a response to the Transcendentalist Challenge to naturalized phenomenology by sketching the contours of what I called the Integrationist View. Such a view, I have argued, is required if we want to not only *allow* for a relationship of mutual enlightenment between phenomenology and science (as MT does) but also *make sense* of it. The key to this view is the conception of consciousness as a structure of behavior, ontologically prior to the distinctions between objectivity and subjectivity and third- and first-person perspectives. As we have seen, this "ontology of structure" is in Merleau-Ponty's philosophy motivated not only by the transcendentally naive perspective of observers of behavior but also equally through an internal critique of the transcendental perspective itself. In this way, we arrive at a view of the transcendental as not essentially separate from the natural, but rather as organizational norms of contingent, living nature that are best illuminated through a dialectical exchange between phenomenological and scientific approaches.

It might be objected that the end result has abandoned transcendental philosophy altogether. Given a certain conception of "transcendental," that is probably true. However, given the internal critique of transcendentalism involved in IV together with the historical fluidity of transcendental philosophy, I believe the label can be kept if desired.

In the last section, I made an attempt to show a concrete example of phenomenology-science integration, arguing that Merleau-Ponty's engagement with the Schneider case has the potential to both inform the scientific understanding of pathology and to be a moment in radical reflection's uncovering of its own conditions. Thus, the integrationist view finally went from abstract articulation toward a concrete *sense*, for, as Merleau-Ponty says in the last page of the *Phenomenology*, "philosophy actualizes itself by destroying itself as an isolated philosophy" (2012, p. 483).

AUTHOR CONTRIBUTIONS

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

ACKNOWLEDGMENTS

Thanks to Ståle Finke, Jonathan Knowles, Sigurd Hverven, Mads Dengsø, Michael Amundsen, Maia Vige Helle, the VERP colloquium at NTNU, and the two reviewers for very helpful comments on earlier drafts of this paper. I'm also grateful to the Center for Subjectivity Research at the University of Copenhagen for providing the stimulating environment in which the initial idea and sketch of the article were formed.

REFERENCES

Ameriks, K. (2003). "Kant's transcendental deduction as a regressive Argument," in *Interpreting Kant's Critiques*, ed. K. Ameriks (Oxford: Oxford University Press), 51–66.

Bhaskar, R. (1978). *A Realist Theory of Science*. Sussex: The Harvester Press.

Canguillhem, G. (1978). *On the Normal and the Pathological*. London: D. Reidel Publishing Company. R. Fawcett.

Carr, D. (1999). *The Paradox of Subjectivity: The Self in the Transcendental Tradition*. Oxford: Oxford University Press.

De Preester, H. (2002). Naturalizing husserlian phenomenology: an introduction. *Psychoanal. Perspect.* 20, 633–647.

Farah, M. (2004). *Visual Agnosia*. Cambridge, MA: MIT.

Fernandez, A. V. (2015). Contaminating the transcendental. *J. Speculat. Philos.* 23, 291–301.

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

Gallagher, S. (2017). *Enactivist Interventions: Rethinking the Mind*. Oxford: Oxford University Press.

Gardner, S. (2015). "Merleau-Ponty's transcendental theory of perception," in *The Transcendental Turn*, eds S. Gardner and M. Gist (Oxford: Oxford University Press).

Goldstein, K., and Gelb, A. (1918). Psychologische analysen hirnpathologischer fälle auf grund von untersuchungen hirnverletzer. *Z. Gesamte Neurol. Psychiatrie* 41, 1–142. doi: 10.1007/bf02874477

Habermas, J. (1991). "Philosophy as stand-in and interpreter," in *Moral Consciousness and Communicative Action: Studies in Contemporary German Social Thought*, by Jürgen Habermas, eds C. Lenhardt and S. W. Nicholsen (Cambridge: MIT Press), 1–20.

Henrich, D. (1969). The proof-structure of Kant's transcendental deduction. *Rev. Metaphys.* 4, 640–659.

Husserl, E. (1970). *The Crisis of the European Sciences and Transcendental Phenomenology*. Evanston: Northwestern University Press. Translated by David Carr.

Inkpin, A. (2017). Was Merleau-Ponty a 'Transcendental' Phenomenologist? *Cont. Philos. Rev.* 50, 27–47. doi: 10.1007/s11007-016-9394-0

Jansen, J. (2015). Transcendental philosophy and the problem of necessity in a contingent world. metodo. *Int. Stud. Phenomenol. Philos.* 1, 47–80. doi: 10.19079/metodo.s1.1.47

Jensen, R. T. (2009). Motor intentionality and the case of Schneider. *Phenomenol. Cogn. Sci.* 8, 371–388. doi: 10.1007/s11097-009-9122-x

Kant, I. (2007). *Critique of Pure Reason*, eds M. Weigelt and M. Müller (London: Penguin Books Ltd).

Kee, H. (2020). Phenomenological reduction in Merleau-Ponty's the structure of behavior: an alternative approach to the naturalization of phenomenology. *Eur. J. Philos.* 28, 15–32. doi: 10.1111/ejop.12452

McDowell, J. (1994). *Mind and World*. Cambridge: Harvard University Press.

Merleau-Ponty, M. (1963). *The Structure of Behavior*. Pittsburgh: Duquesne University Press.

Merleau-Ponty, M. (1964). "Phenomenology and the sciences of man," in *The Primacy of Perception*, ed. J. Wild (Evanston: Northwestern University Press), 43–95.

Merleau-Ponty, M. (1968). *The Visible and the Invisible*, ed. C. Lefort (Evanston: Northwestern University Press). Translated by Alphonso Lingis.

Merleau-Ponty, M. (2012). *Phenomenology of Perception*. New York, NY: Routledge.

Moran, D. (2013). 'Let's look at it objectively': why phenomenology cannot be naturalized. *R. Inst. Philos. Suppl.* 72, 89–115. doi: 10.1017/s1358246113000064

Morris, D. (2018). *Merleau-Ponty's Developmental Ontology*. Evanston: Northwestern University Press.

Reynolds, J. (2018). *Phenomenology, Naturalism and Science: A Hybrid and Heretical Proposal*. New York, NY: Routledge.

Roux, J. M. (2013). Naturalism and transcendentalism: the ubiquity of idealism. *Metodo. Int. Stud. Phenomenol. Philos.* 1, 197–213.

Roy, J. M., Petitot, J., Pachoud, B., and Varela, F. J. (1999). "Beyond the Gap: an introduction to naturalizing phenomenology," in *Naturalizing Phenomenology: Issues in Contemporary Phenomenology and Cognitive Science*, eds J. Petitot, F. J. Varela, B. Pachoud, and J. M. Roy (Stanford: Stanford University Press), 1–80.

Stern, R. (2000). *Transcendental Arguments and Scepticism: Answering the Question of Justification*. Oxford: Oxford University Press.

Strawson, P. F. (1966). *The Bounds of Sense: An Essay on Kant's Critique of Pure Reason*. London: Routledge.

Taylor, C. (1997). "Overcoming epistemology," In *Philosophical Arguments*, ed C. Taylor (Cambridge: Harvard University Press).

Thompson, E. (2007). *Mind in Life: Biology, Phenomenology, and the Sciences of Mind*. Cambridge: Harvard University Press.

Toadvine, T. (2009). *Merleau-Ponty's Philosophy of Nature*. Evanston: Northwestern University Press.

Vanzago, L. (2012). Naturalizing Phenomenology and the nature of phenomena. On varela, petitot, and merleau-ponty. *Chiasmi Int.* 14, 131–142. doi: 10.5840/chiasmi20121413

Weber, A., and Francisco, J. V. (2002). Life after kant: natural purposes and the autopoietic foundations of biological individuality. *Phenomenol. Cogn. Sci.* 1, 97–125.

Zahavi, D. (1996). Husserl's intersubjective transformation of transcendental philosophy. *J. Br. Soc. Phenomenol.* 1996, 228–245.

Zahavi, D. (2004). Phenomenology and the project of naturalization. *Phenomenol. Cogn. Sci.* 2004, 331–347. doi: 10.1023/b:phen.0000048935.94012.4e

Zahavi, D. (2008). "Phenomenology," in *The Routledge Companion to Twentieth Century Philosophy*, ed. D. Moran (London: Routledge), 661–692.

Zahavi, D. (2010). "Naturalized phenomenology," in *Handbook of Phenomenology and Cognitive Science*, eds S. Gallagher and D. Schmicking (London: Springer), 3–19.

Zahavi, D. (2011). Mutual enlightenment and transcendental thought. *J. Conscious. Stud.* 2011, 169–175.

Zahavi, D. (2013). Naturalized phenomenology: a desideratum or a category mistake? *R. Inst. Philos. Suppl.* 2013, 23–42. doi: 10.1017/s1358246113000039

Zahavi, D. (2017). *Husserl's Legacy*. Oxford: Oxford University 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

HYPOTHESIS AND THEORY
published: 21 July 2020
doi: 10.3389/fpsyg.2020.01680

The Hitchhiker's Guide to Neurophenomenology – The Case of Studying Self Boundaries With Meditators

Aviva Berkovich-Ohana[1,2,3,4], Yair Dor-Ziderman[1,2,3], Fynn-Mathis Trautwein[3,5], Yoav Schweitzer[1,2,3], Ohad Nave[6], Stephen Fulder[7] and Yochai Ataria[8]*

[1] Department of Learning, Instruction and Teacher Education, Faculty of Education, University of Haifa, Haifa, Israel, [2] Department of Counseling and Human Development, Faculty of Education, University of Haifa, Haifa, Israel, [3] Edmond J. Safra Brain Research Center, University of Haifa, Haifa, Israel, [4] The Integrated Brain and Behavior Research Center (IBBRC), University of Haifa, Haifa, Israel, [5] Department of Psychosomatic Medicine and Psychotherapy, Medical Center – University of Freiburg, Freiburg im Breisgau, Germany, [6] Department of Cognitive Sciences, The Hebrew University of Jerusalem, Jerusalem, Israel, [7] The Israel Insight Society (Tovana), Karmiel, Israel, [8] Department of Psychology, Tel-Hai Academic College, Tel Hai, Israel

This paper is a practical guide to neurophenomenology. Varela's neurophenomenological research program (NRP) aspires to bridge the gap between, and integrate, first-person (1P) and third-person (3P) approaches to understanding the mind. It does so by suggesting a methodological framework allowing these two irreducible phenomenal domains to relate and reciprocally support the investigation of one another. While highly appealing theoretically, neurophenomenology invites researchers to a challenging methodological endeavor. Based on our experience with empirical neurophenomenological implementation, we offer practical clarifications and insights learnt along the way. In the first part of the paper, we outline the theoretical principles of the NRP and briefly present the field of 1P research. We speak to the importance of phenomenological training and outline the utility of cooperating with meditators as skilled participants. We suggest that 1P accounts of subjective experience can be placed on a complexity continuum ranging between thick and thin phenomenology, highlighting the tension and trade-off inherent to the neurophenomenological attempt to naturalize phenomenology. We then outline a typology of bridges, which create mutual constraints between 1P and 3P approaches, and argue for the utility of alternating between the bridges depending on the available experimental resources, domain of interest and level of sought articulation. In the second part of the paper, we demonstrate how the theory can be put into practice by describing a decade of neurophenomenological studies investigating the sense of self with increasing focus on its embodied, and minimal, aspects. These aspects are accessed via the dissolution of the sense-of-boundaries, shedding new light on the multi-dimensionality and flexibility of embodied selfhood. We emphasize the evolving neurophenomenological dialogue, showing how consecutive studies, placed differently on the thin-to-thick 1P continuum, advance the research project by using the bridging principles appropriate for each stage.

Keywords: neurophenomenology, sense of self, embodied self, self boundaries, meditation, magnetoencephalography (MEG)

OPEN ACCESS

Edited by:
Wenjing Cai,
Shanghai Jiao Tong University, China

Reviewed by:
Urban Kordes,
University of Ljubljana, Slovenia
Sucharit Katyal,
UC Davis Medical Center,
United States

****Correspondence:***
Aviva Berkovich-Ohana
avivabo@edu.haifa.ac.il

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

Received: *27 February 2020*
Accepted: *19 June 2020*
Published: *21 July 2020*

Citation:
Berkovich-Ohana A, Dor-Ziderman Y, Trautwein F-M, Schweitzer Y, Nave O, Fulder S and Ataria Y (2020) The Hitchhiker's Guide to Neurophenomenology – The Case of Studying Self Boundaries With Meditators. Front. Psychol. 11:1680. doi: 10.3389/fpsyg.2020.01680

INTRODUCTION

Toward the end of the last millennium, Francisco Varela, the Chilean biologist, philosopher, and neuroscientist, put forth an ambitious proposal, the neurophenomenological research program (NRP). The NRP attempts at paving a methodological path for bridging the 'explanatory gap' in our understanding of how to integrate first-person (1P) phenomenological and third-person (3P) physiological features of the mind. Instead of ineffective attempts to close the conceptual gap between subjective experience and physical matter, also referred to as the 'Hard Problem of Consciousness' (Chalmers, 1995), the NRP suggests reframing the gap. To this end, it defines a methodological strategy for integrating phenomenological and neurobiological accounts under one research program, or in Varela's words, for creating "meaningful bridges between two irreducible phenomenal domains" (Varela, 1996, p. 330). The NRP has been described as "a quest to marry modern cognitive science and a disciplined approach to human experience" (Ibid. p. 340) and it is this harmonious relationship which we pursue.

With more than two decades since its introduction, numerous theoretical studies have illuminated the importance, ingenuity and potential of the NRP, yet not as many have succeeded in actually implementing it as an empirical methodology (for a short review see **Tables 1**, **2**; for other reviews, see Lutz and Thompson, 2003; Bayne, 2004; Thompson et al., 2005; Lutz, 2006, 2007; Bitbol, 2012; Bockelman et al., 2013; Ataria, 2017; Berkovich-Ohana, 2017). Clearly, NRP's strong philosophical appeal is merely the grounds from which a scientific paradigm may grow and mature to address the highly ambitious and demanding challenge at stake – "an open-ended quest for resonant passages between human experience and cognitive science" (Varela, 1996, p. 346). If the challenge is to be met on Varela's terms, it can only be done based on rigorous empirical practice. It is the goal of this paper to support those who wish to embark on this journey and implement the NRP by highlighting some of the important lessons learnt in neurophenomenological studies done in recent years, including our own.

In the first part of the paper, we present Varela's NRP while specifically focusing on various issues regarding the execution of an empirical phenomenological investigation. We point toward an inherent tension within the NRP concerning the challenge of naturalization in the face of the complexity and intricacy of 1P data. In this regard, we offer a diverse typology of bridges which exemplifies the concept of 'mutual constraints,' and argue for the need to gradually and interchangeably weave them through the developmental stages of an evolving research program.

The second part of this paper demonstrates our own decade-long implementation of the NRP focused on breaking new ground in the scientific understanding of self consciousness, with particular interest in alterations in the sense of embodied self and the minimal self in meditative experience. We found it helpful to repeatedly circulate between different forms of phenomenological inquiry and a variety of cognitive and neuroscientific tools, and argue that it is the ongoing development of a dialogue between these two perspectives that enables novel insights. We close the paper by presenting our current aims of advancing the understanding of self consciousness by employing an ongoing mature and pragmatic neurophenomenological study of the sense of self boundaries and their dissolution.

VARELA'S NEUROPHENOMENOLOGICAL RESEARCH PROJECT (NRP)

The Archimedean point of the NRP is acknowledging the irreducible nature of conscious experience: "Lived experience is where we start from and where all must link back to, like a guiding thread" (Varela, 1996, p. 334). Stemming from the phenomenological tradition (see **Supplementary Appendix 1** for an outline of its historical roots), this notion has far-reaching implications for how we conceptualize nature and our place as embodied cognitive agents within it. It reminds us of the ineradicability of our own standpoint as humans (or cognitive scientists) and motivates a search for an understanding of the co-determination of mind and world as a middle way between the dead-ends of realism or idealism (Varela et al., 1991). Theoretically, this aim is pursued within the enactive approach, which is increasingly becoming a stronghold in cognitive science driving manifold research agendas (see Stewart et al., 2010; Vörös et al., 2016 and Newen et al., 2018 for overviews). Given their common theoretical starting point, neurophenomenology can be considered as a promising methodology of enactivism (Vörös et al., 2016). It seeks to provide a pragmatic methodological framework in which cognitive neuroscience can rigorously integrate a disciplined examination of conscious experience. This notion stresses the necessity of acquiring refined and reliable 1P descriptions in order to advance toward "a model that can account for both the phenomenology and neurobiology of consciousness in an integrated and coherent way" (Thompson et al., 2005, p. 87). The emphasis here is on forming constant circulation and dialogue between these two domains of inquiry which would allow an exploration of "the bridges, challenges, insights and contradictions between them" (Varela, 1996, p. 343).

Inspired by the phenomenological tradition, the call for a systematic exploration of lived experience put forward by the NRP has received considerable attention in recent years. In essence, such exploration is grounded in a set of practices that generally allow subjects to increase their sensitivity to their moment to moment experience (Varela and Shear, 1999; Petitmengin, 2011). Stemming from Edmund Husserl's concepts of the phenomenological reduction and the epoché (Husserl, 1970; Bitbol, 2019a; for elaboration, see **Supplementary Appendix 1**), there are currently various first person (1P) and second person (2P)[1] methods that promote a shift in attitude from the natural theory-laden absorption with the contents of one's experience, to an awareness of the various affective, attentional and structural features of experience (discussed

[1]This therefore is a method enabling the gathering of "first person" data, i.e., data that express the viewpoint of the subject himself, in the grammatical form "I. . .". But as these data have been gathered through another person (a "You"), it has been dubbed a "second person" method (Petitmengin, 2006, p. 231).

TABLE 1 | A list of papers relating to neurophenomenology (NP) on the theoretical level.

Theoretical perspectives	Papers
What is NP	Varela, 1996; Lutz, 2002, 2007; Lutz and Thompson, 2003; Rudrauf et al., 2003; Thompson, 2006
'Front-loading' phenomenology into experimental design	Gallagher and Sørensen, 2006
Neurofeedback useful for NP	Bagdasaryan and Le Van Quyen, 2013
Hypnosis useful for NP	Lifshitz et al., 2013
Astronaut simulation and challenges of NP	Bockelman et al., 2013
Dreams and challenges of NP	Solomonova et al., 2014
Kantian perspective on NP	Khachouf et al., 2013
NP in the context of in dyadic movement	Stuart, 2013
Application of NP in affective neuroscience	Colombetti, 2013
NP usefulness in psychology	Gordon, 2013
NP usefulness in ECoG	Petitmengin and Lachaux, 2013
NP justification in practice	Strle, 2013
Philosophical attack and justification	Bitbol and Antonova, 2016; Kirchhoff and Hutto, 2016
Advantage of training scientists in contemplation	Desbordes and Negi, 2013
NP applied to time consciousness	Varela, 1999
NP usefulness in pain	Price et al., 2002
Combining descriptive experience sampling with 3P	Hurlburt et al., 2017
Combining 2P methods	Froese et al., 2011; Olivares et al., 2015
Philosophical roots of NP and enaction	Vörös et al., 2016
An externalist extension of NP	Beaton, 2013
Adding a cardiac-affective dimension to NP	Depraz and Desmidt, 2018
Distinguishing mild vs. radical NP	Petitmengin, 2017
The feasibility of NP	Ataria, 2017
Neurofeedback useful for NP	Bielas and Michalczyk, 2017
NP usefulness to understand trauma	Ataria et al., 2019
Application of NP to microdreams	Nielsen, 2017
Phenomenological matrix of mindfulness	Lutz et al., 2015
Phenomenologically constrained neurocomputational model of the self	Williford et al., 2018
Brain dynamics from the perspective of NP	Fazelpour and Thompson, 2015
NP of surprise	Bitbol, 2019b
Neuroscience and inner experience	Price and Barrell, 2012

further in section "Investigating Lived Experience – From Thin to Thick Phenomenology"). In other words, instead of focusing on the 'what,' subjects are encouraged to bracket assumptions and presuppositions *about* their experience and become aware of the 'how' of experience, that is, the subjective mode of appearance and the dynamic intentional acts involved in the flow of experience (Petitmengin, 2006; Bitbol, 2019a). This dimension of experience is often termed by phenomenologists as pre-reflective awareness pointing toward the tacit, direct and non-inferential awareness of one's experience as it is lived through, prior to any second-order reflection *on* experience (Nagel, 1974; Merleau-Ponty, 1996; Zahavi, 2002).

1P accounts have a vital role to play, along neuroscience and physiology, in beginning to bridge the explanatory gap *a la* Varela. The working hypothesis of the NRP is that "phenomenological accounts of the structure of experience and their counterparts in cognitive science relate to each other through reciprocal constraints" (Varela, 1996, p. 343). It is this type of circulation between the two perspectives which is at the focus of this paper and is described in more detail in Section "Building Bridges Between Phenomenology and Physiology – Mutual Constraints." By reciprocal constraints Varela means both using neuroscientific accounts to illuminate previously unnoticed aspects of mental experience, and on the other hand, guiding the empirical questions, analysis and interpretation of neurobiological findings in light of the phenomenal invariants of the mental experience. 1P data generated from the phenomenology of mental processes "can provide additional, valid information about externally uncontrollable aspects of mental activity, and this information can be used to detect significant patterns of dynamic activity at the neural level" (Thompson et al., 2005, pp. 45–46). Thus on a methodological level, the NRP suggests explicitly and rigorously incorporating phenomenological investigation into experimental setup and design.

Investigating Lived Experience – From Thin to Thick Phenomenology

The question of the importance, validity and place of the investigation of lived experience within science has seen many diverse conceptualizations throughout history. It is obviously

TABLE 2 | A list of papers employing neurophenomenological empirical paradigms.

1P/2P	3P	Papers
Attentional state	EEG	Lutz et al., 2002
	Behavior	Van den Bussche et al., 2013; Zanesco et al., 2013
Epilepsy	EEG	Le Van Quyen and Petitmengin, 2002; Petitmengin et al., 2006, 2007
Sense-of-self (time/space/boundary/identification)	MEG	Berkovich-Ohana et al., 2013a; Dor-Ziderman et al., 2013; Ataria et al., 2015; Dor-Ziderman et al., 2016
	fMRI	Garrison et al., 2013
Hypnosis	EEG	Cardeña et al., 2013
Astronaut simulation	EEG/fNIRS	Reinerman-Jones et al., 2013
Mind-wandering	fMRI	Christoff et al., 2009; Allen et al., 2013
Emotion	fMRI	Northoff and Heinzel, 2006
Yoga/attention/emotion	HRV	Mackenzie et al., 2014
Neuro inspires pheno		De Haan et al., 2013; Valenzuela Moguillansky et al., 2013
Resting state	fMRI/EEG	Diaz et al., 2013
Intention to act	EEG	Guggisberg and Mottaz, 2013; Jo et al., 2014, 2015
Dreaming	PET/fMRI	Fox et al., 2013
Memory	HRV/GSR	McCall et al., 2015
Pain under hypnosis	PET	Rainville et al., 1997
Language/auditory processing	fMRI	Kühn et al., 2014; Hurlburt et al., 2016
Descriptive experience sampling	Verbal behavior	Hurlburt et al., 2002
Approach-avoidance	Behavior	Baquedano and Fabar, 2017
Surprise in depression	ECG	Depraz et al., 2017
Meditative state	Behavior	Abdoun et al., 2019
Psychedelic state	EEG	Timmermann et al., 2019

EEG, electroencephalography; MEG, magnetoencephalography; fMRI, functional magnetic resonance imaging; HRV, heart rate variability; PET, positron emission tomography; GSR, galvanic skin response; fNIRS, functional near-infrared spectroscopy.

beyond the scope of this paper to address the often-encountered notion (in cognitive neuroscience circles in particular) that subjective experience cannot be fully studied using the scientific method (but see Varela and Shear, 1999; Depraz et al., 2003; Froese et al., 2011; Olivares et al., 2015; Kordeš and Demšar, 2018). The aim of this section is to briefly present the state of the art in the developing field of 1P research, highlighting the demands, obstacles and practical implications imposed on neurophenomenological studies. We then present the concepts of 'thin' and 'thick' phenomenology as an organizational tool for the different levels of depth which specify 1P data in the context of the NRP.

As lived experience is a constantly changing, multi-layered and highly complex flux, its exploration is challenging. It requires employing a certain mental gesture of reflection toward one's own experience that differs considerably from casual 'introspection.' Though formulated in various ways, it is often argued by 1P researchers that "becoming aware of lived experience is a skill that can and should be learned and practiced" (Froese et al., 2011, p. 254). Indeed, a rather demanding element in Varela's radical vision is that cognitive science students interested in mental experience "must inescapably attain a level of mastery in phenomenological examination" through sustained training (Varela, 1996, p. 347). While such training may intuitively seem valuable to researchers' own understanding of the phenomena they are investigating (Jack and Roepstorff, 2002; Gallagher and Zahavi, 2008), in the context of experimental neurophenomenological studies this effort is often centered on the rigorous acquisition of reliable 1P data from study participants. One suggested way to acquire such data is by using phenomenologically trained subjects (discussed in section "Cooperation With Meditators As Skilled Participants"), while another is based on interview techniques collectively known as second-person (2P) methods. Thus, first and second -person methods serve as supports for the examination of lived experience. 1P methods cultivate the capacity for sustained awareness, helping subjects gain access to aspects of their experience that are lived through but mostly remain unnoticed (Petitmengin, 2006; Froese et al., 2011); while the disciplined nature of 2P methodologies enables the systematic gathering of reliable phenomenological reports which can then be incorporated in various ways into neuroscientific research (elaborated in the following sections).

Recently, several interview techniques designed for detailed and careful description of subjective experience have been developed (for an overview, see Froese et al., 2011; Olivares et al., 2015). Mediated by the guidance of a skillful interviewer, they allow "a person, who may not even have been trained, to become aware of his or her subjective experience, and describe it with great precision" (Petitmengin, 2006, p. 334). Two notable 2P methods, Micro-phenomenology (Petitmengin, 2006; also known as 'Elicitation Interview,' Vermersch, 2009) and Descriptive Experience Sampling ('DES,' Hurlburt and Akhter, 2006), are based on retrospective examination of past experiences framed and guided by an empathetically tuned phenomenological investigator. In the context of neurophenomenology, these

methods systematize the phenomenological research procedure, thus serving as valuable tools for performative coherence and scientific rigor. While the issue of validity of such procedures remains widely contentious (Dennett, 1991; Searle, 1992), progress has been made in recent years in cultivating pragmatics that help reduce the influence of bias, increase authenticity and evaluate reliability (Petitmengin, 2006; Olivares et al., 2015; Kordeš et al., 2019; Petitmengin et al., 2019a; Valenzuela-Moguillansky and Vásquez-Rosati, 2019). Following other scientific domains, data obtained through phenomenological inquiry is not taken at face value as infallible but examined, interpreted, analyzed for invariant structures and generalized in various ways. Eventually, the status of 1P accounts is not determined by their facticity, but evaluated through procedural standardization, potential replication of its findings and intersubjective validation with other first, second and third -person methods. As Varela suggests:

> "The usual opposition of first-person vs. third-person accounts is misleading. It makes us forget that so-called third-person, objective accounts are done by a community of concrete people who are embodied in their social and natural world as much as first-person accounts [. . .]. The line of separation — between rigor and lack of it — is not to be drawn between first and third person accounts, but determined rather by whether there is a clear methodological ground leading to a communal validation and shared knowledge" (Varela, 1996, p. 340).

It is this methodological ground which we aim to advance. As practitioners of integrated 1P and 3P research, we have gained some experience in the pragmatics of neurophenomenology. Thus, we are far from naively suggesting that 1P methods can be added into neuroscience without reflection, but rather hope to further illuminate some practical considerations that address inherent challenges.

Embracing the pragmatic spirit of the NRP, we suggest framing the diversity of possible forms of 1P data on a continuum of complexity (**Figure 1**). On one pole there is data obtained through in-depth phenomenological interview methods (Bitbol and Petitmengin, 2016a; Petitmengin et al., 2017, 2019b; Ollagnier-Beldame and Cazemajou, 2019) characterized by highly refined, detailed and dynamic accounts of singular subjective experiences (such as data gathered through micro-phenomenology). We call this thick phenomenology, to denote the high complexity of 1P data. On the other pole are thin methods (e.g., self-reports and questionnaires), acquiring data which provides information relevant to subjective experience, but limited due to its reductive nature (using pre-defined rigid phenomenal invariants) which also makes it prone to biases (Polkinghorne, 1989; Hurlburt and Heavey, 2015). We call it thin, to denote the low level of complexity of this kind of 1P data. In between these two poles, various methods can be placed, for example structured interviews and Descriptive Experience Sampling (DES, Hurlburt and Akhter, 2006)[2].

The continuum directly relates to the possibility of naturalizing subjective experience, which in this specific respect refers to the prospect of integrating 1P account within an experimental neuroscientific setup[3]. At the thinner end of the continuum, data can be acquired rapidly, repeatedly and uniformly, facilitating intersubjective and cross-situational generalization. This is often relevant for its integration with neural measures because these typically require a larger number of sampled timepoints and individuals in order to yield reliable results. Furthermore, thin 1P data is easier to formalize and often quantifiable and thus more suitable for guidance of 3P data analysis, as well as easier accessible for scientific dialogue and cross-validation. Conversely, it is often limiting, unreliable and tainted by artifacts as it fails to address the multi-layered intricacy and dynamics of lived experience and to bracket assumptions and presuppositions of respondents (Hurlburt and Heavey, 2015; Bitbol and Petitmengin, 2016b).

Thicker methods of investigation result in much more refined accounts of experience, potentially amounting to more authentic descriptions of experience sensitive to its multi-dimensional and diachronic nature. Their skillful execution, analysis and formalization are thus far more cumbersome and require meticulous effort. They are, on the one hand, less accessible for generalization and quantification, but on the other hand, their high resolution may permit the bridging of experiential and neural microdynamics (as proposed by Petitmengin and Lachaux, 2013). Refined accounts of experience are also more receptive to novel insights which can guide future research. Overall, the naturalization of these methods through efficient integration within an experimental setup remains a challenging endeavor.

Rather than advocating the use of one method over the other, we suggest that this methodological trade-off is essential to the open-ended circulation envisioned by the NRP. As elaborated in the next section, we found it helpful in our own studies to alternate between different methods depending on factors regarding the available neuroscientific and phenomenological resources, as well as considering the current level of understanding of the studied phenomena. Such an integrative mixed-methods approach is useful in the triangulation of acquired data, and enjoys the advantage of using precise techniques that access specific and distinct dimensions of lived experience (further discussed in section "Building Bridges Between Phenomenology and Physiology – Mutual Constraints").

Cooperation With Meditators as Skilled Participants

As 1P data is fundamental for the NRP, scientific cooperation with "phenomenologically trained subjects" has been suggested

[2]In DES, pre-trained participants are 'beeped' randomly, thus ordered to record (verbally or in writing) their inner experience in that particular moment. DES would be placed in between the two poles, as the resolution of the lived experience obtained in this method is lower, in part, in order to minimize bias of reconstruction. Importantly, it is used to inspect random, rather than specifically meaningful instants of experience.

[3]The issue of naturalization of phenomenology clearly exceeds the specific context in which it is discussed here. It is further explored through the typology of bridges provided in Section "Building Bridges Between Phenomenology and Physiology – Mutual Constraints." For theoretical accounts of naturalization of phenomenology, see Petitot et al. (1999), De Preester (2002), Overgaard (2004), Zahavi (2004), Gallagher (2012).

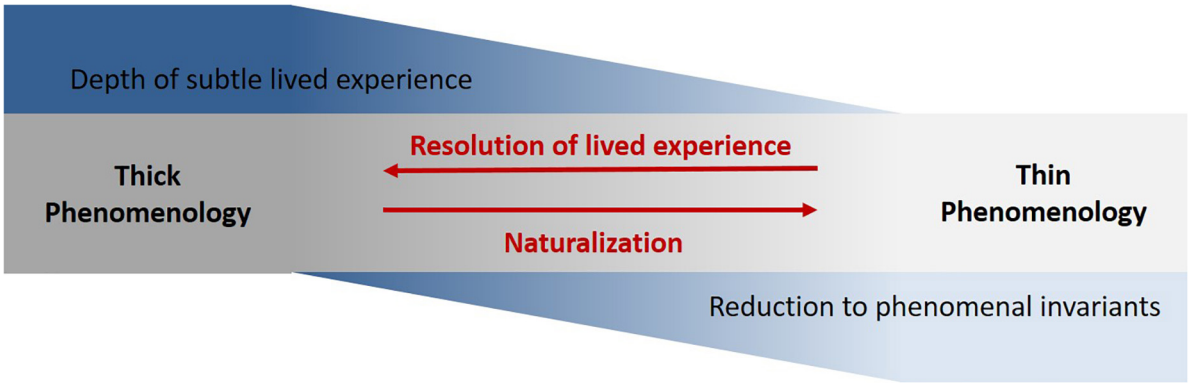


FIGURE 1 | Tension and trade-off between phenomenological thickness and naturalization. The thick and thin phenomenology extremes form a continuum of the depth of 1P data acquisition.

to be useful (Thompson et al., 2005, p. 45). While such training is effortful and time-consuming, senior meditation practitioners have been proposed by several authors as pre-trained subjects suitable for such inquiry, based on the similarity between the epoché and certain meditative techniques (Varela et al., 1991; Varela and Shear, 1999; Depraz et al., 2003; Bitbol, 2019a; Depraz, 2019; Kordeš et al., 2019; Vörös, 2019). For example, Bitbol (2019a) argues that the common characterization of mindfulness, "namely as concentrating on the present moment and staying non-judgmental, captures a fundamental feature of the epochè: not only the suspension of elaborate judgments, but even before that, the suspension of the semantic function of both mental and verbal activities, that tends to expel us from our present." (Ibid., p. 136). Similarly, Varela suggested that certain forms of meditation from the Buddhist tradition can be conceived of as epoché:

> The exercise of samatha, best translated as mindfulness meditation practice, is based on an examination of the nature of our mind, (and hence of the origin of habitual patterns) by paying meticulous attention to every moment of appearance. In other words, using the activity of mind to go beyond mind, looking at the givenness of experience with a fresh, inquiring glance (Varela, 2000, p. 5).

In addition to the cultivation of the ability to suspend judgments and maintain a fresh conceptual-free perspective, other potential assets trained meditators bring to phenomenology are the ability to volitionally reproduce specific features of experience as cultivated in a given meditative practice (Lutz, 2002; Lutz et al., 2007; Desbordes and Negi, 2013), as well as the ability to stay with the experience being studied, i.e., reduce getting 'lost in thought' and mind-wandering as is typical in untrained subjects (Mrazek et al., 2013).

While a number of authors have claimed that the cultivation of mindfulness improves the quality of introspection, it has not yet been sufficiently empirically established, yet some supportive evidence is available. For example, meditative practice was shown to correlate with the predictive accuracy of self-reports regarding behavior (Abdoun et al., 2019), neuroanatomy (Fox et al., 2012) and peripheral physiology (Parkin et al., 2014). The latter was also confirmed in a recent meta-analysis indicating a small but statistically significant positive relationship between mindfulness and objective measures of body awareness (Treves et al., 2019). This issue is still a matter of debate due to the complexity of its assessment.

It is important to mention that alongside substantial benefits, there are also concerns and drawbacks of harnessing qualified meditators as "phenomenologically trained subjects." First, it might be hard for some meditators to examine their experience beyond the elements prescribed by their practice school, or to let go of the conceptualizations of their respective traditions (Kordeš et al., 2019). Second, the goal of the observation is different: while Buddhist practices aim at alleviating human suffering, the central motivation of phenomenology is knowledge (Bitbol, 2019a). Thus, in the long-term it might be beneficial for science to establish its own paradigms of contemplative (phenomenological) training as envisioned by Varela.

With this in mind, cooperation with meditators has proved to be useful for the NRP not only for their alleged familiarity with the phenomenological attitude, but also for their enhanced sensitivity to subtle aspects of their experience (see Jo et al., 2014), as well as their capacity to volitionally control and stably maintain specific conscious (and neural) states (for some empirical evidence, see Garrison et al., 2013, 2015). This last ability increases the signal-to-noise ratio and renders these features scientifically tractable. It is the combination of these elements borne by meditative practice that supports the scientific exploration of subtle aspects of consciousness as exhibited in our own research and that of others (Berkovich-Ohana et al., 2013a; Dor-Ziderman et al., 2013, 2016; Zanesco et al., 2013; Jo et al., 2014; Ataria et al., 2015; Lutz et al., 2015; Abdoun et al., 2019).

Building Bridges Between Phenomenology and Physiology – Mutual Constraints

The working hypothesis of neurophenomenology is that phenomenological accounts of the structure of experience and

their counterparts in cognitive science relate to each other through reciprocal constraints. According to Varela (1996):

> It is quite easy to see how scientific accounts illuminate mental experience, but the reciprocal direction, from experience toward science, is what is typically ignored. What do phenomenological accounts provide? At least two main aspects of the larger picture. First, without them the firsthand quality of experience vanishes, or it becomes a mysterious riddle. Second, structural accounts provide constraints on empirical observations (p. 343).

There are different ways to create bridges between the two irreducible phenomenal domains of experience and physiology. In this section, we offer a typology of previously proposed bridges, give an example to each of them, and schematically illustrate their directionality in **Figure 2**.

Bridge A: Front-Loading Phenomenological Insights Into Experimental Design

Gallagher (2003) and Gallagher and Sørensen (2006) suggested as a bridge between phenomenology and physiology the 'front-loading' of phenomenological insights onto experimental design. In other words, to design experiments informed by phenomenological insights – developed in independently conducted phenomenological analyses, or from previous neurophenomenological experiments. Such an approach was successfully implemented in a series of neuroimaging studies on ownership and agency (e.g., Chaminade and Decety, 2002; Farrer and Frith, 2002), which relied on previously generated phenomenological insights (Gallagher, 2000), rather than implementing 1P measures in the studies themselves (detailed in Gallagher and Sørensen, 2006). Additionally, the researcher's direct access to his own lived experience inevitably influences the design and interpretation of the results. Rather than sweeping it under the carpet, such influences are here acknowledged, reflected upon and refined so that they could enhance the quality of the research[4] (Jack and Roepstorff, 2002; Gallagher and Zahavi, 2008). In **Figure 2**, 'front-loading' phenomenological insights onto experimental design is illustrated as an arrow from 1P to 3P, signifying the use of 1P to design 3P studies.

Bridge B: Phenomenological Validation of Neurobiological Accounts

Varela (1996) proposed that disciplined 1P accounts ought to play be an integral element in the validation of a neurobiological proposal, i.e., that any attempt to scientifically explain mind and consciousness must directly relate to the nature of our lived experience (Varela, 1996, pp. 344–345). An adequate theoretical framework is thus needed to characterize neurophysiology in suitable terms that can also address the essential structure and dynamics of experience (further explored in bridge e). Such an approach proved highly productive in the important works of Petitmengin, Navarro, and Le Van Quyen concerning seizure anticipation (Le Van Quyen and Petitmengin, 2002; Petitmengin et al., 2006, 2007). Since 1975, researchers have used EEG analysis for the prediction of seizures, including a preictal state, which is detectable a few seconds before the actual seizure onset on EEG. The group's EEG work showed that seizures do not arise suddenly, but as a transition from the interictal to the ictal state. It also showed that seizures do not correspond to deficient functioning of localized brain areas, but rather to deficient functioning of neural networks. However, the authors write, "the synchrony analysis does not tell us anything about the way this transition and this deficit are (or are not) felt by the patient. It indicates the structure of the cerebral activity, not the nature of the subjective experience that could correspond to it." (Petitmengin et al., 2007, p. 750). Indeed, a phenomenological investigation showed that all nine investigated patients experienced auras (ictal phenomena), while six experienced prodromes (preictal phenomena). Studying the phenomenological dynamics showed that seizures are preceded by (often pre-reflective) symptoms. In **Figure 2**, phenomenological validation of neurobiological accounts is illustrated as an arrow from 1P to 3P, signifying the use of 1P to enhance insight into 3P.

Bridge C: Joint Analyses of 1P and 3P Person Data

Lutz (2002) emphasized joint analyses of 1P and 3P data, which actually means that phenomenal reports guide the analysis of the neuroscientific data[5]. The utility of this approach was illustrated in the seminal work of Lutz (2002) and Lutz et al. (2002), which received much discussion (Jack and Roepstorff, 2002; Gallagher, 2003; Bayne, 2004; Overgaard, 2004). In this study, subjects were trained to recognize stable categories of experience ('phenomenal invariants'), which related to their state of 'preparedness' for the onset of simple visual stimuli presentation. Reports were then grouped into 'phenomenological clusters.' The results provide an outstanding example of the tailored use of 1P data, as the 'phenomenological clusters' were shown to reflect variability in neuronal response which otherwise would have been considered as noise. For example, only one cluster (when subjects reported "steady readiness" to the stimuli, in contrast to either "fragmented readiness" or "spontaneous un-readiness") correlated with high gamma band EEG synchronization over frontal electrodes before stimulation. Importantly, EEG synchronization and behavioral measures of reaction time both correlated with the subjective reports. In **Figure 2**, joint analyses of 1P and 3P data is illustrated by a double-headed arrow connecting 1P and 3P, signifying the mutual use of both domains: grouping 3P for analysis according to 1P categories determined within the same experiment (or the other way around).

Bridge D: Using Physiological Data to Guide Investigation of Subjective Experience

A meaningful bridge could be constraining and guiding 1P data via the physiology itself. An illuminating demonstration of such an approach is the fMRI-neurofeedback study by Garrison et al. (2013) and Van Lutterveld and Brewer (2015). This study assessed how the 1P experience of meditation relates to neural activity in a core region of the default mode network – the

[4]This is especially crucial in the skillful execution of 2P research (Petitmengin et al., 2019a,b).

[5]This is also an example of front-loading, as Lutz et al. (2002) did pilot runs to identify the categories, which were then incorporated into the design.

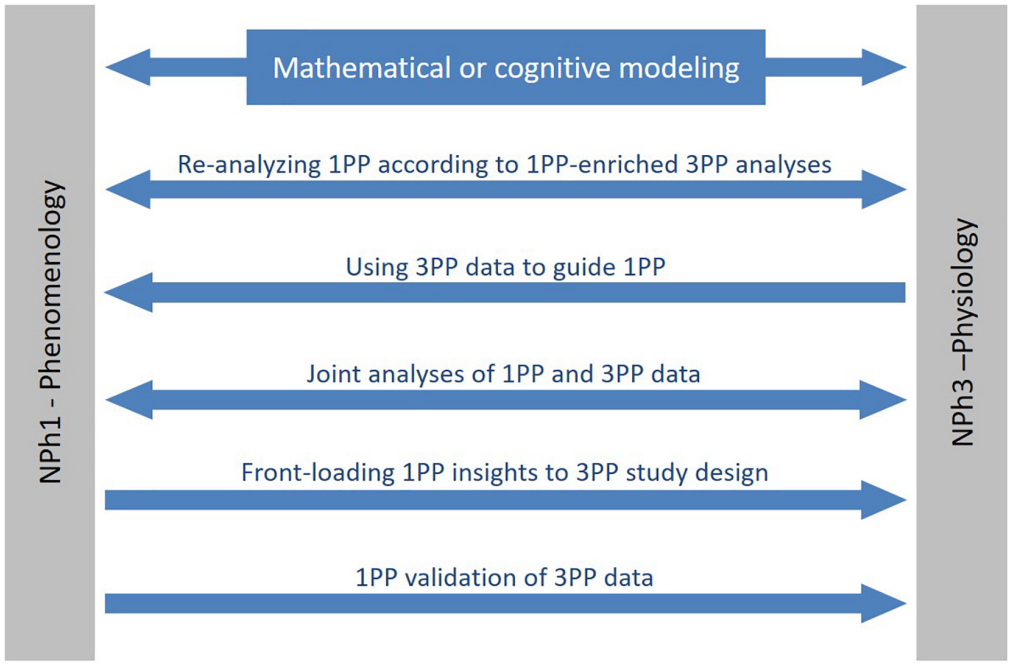


FIGURE 2 | A typology of ways of bridging the irreducible domains of experience and physiology. NPh, neurophenomenology; 1PP, 3PP, first/third person perspective, respectively.

posterior cingulate cortex. Activity in this region was measured and displayed on a screen in real-time, enabling participants to realize how their experience related to changes in the graph. The researchers then analyzed 1P data consisting of meditators' accounts of their subjective experience during runs of a real time fMRI neurofeedback, and 3P data consisting of corresponding PCC activity during the same runs. The results showed that for meditators, subjective experiences corresponding with PCC deactivation related to "undistracted and effortless awareness," while the subjective experiences of "distracted and effortful awareness" corresponded with PCC activation. In **Figure 2**, using 3P data to guide 1P is illustrated as an arrow from 3P to 1P, signifying the use of 3P to constrain and study 1P data.

Bridge E: Re-analyzing the 1P According to the 1P-Enriched 3P Analyses

Another way of creating meaningful bridges between 1P and 3P data is implementing iterative processes, i.e., re-analyzing the 1P according to the 1P-enriched 3P analyses (Lutz, 2002). This is actually the 3rd stage of the formal NRP as proposed by Varela, and hence represents a maturation of one's project. To the best of our knowledge, the only work which implemented this approach is described by Petitmengin et al. (2007) as a "neuro-phenomenological circulation process." In this work focusing on seizure anticipation (detailed in bridge b), the discovery of a new neuro-dynamic structure (the preictal neuro-electric desynchronization) permitted a refinement of the corresponding experiential dynamics (preictal phenomenological symptoms and therapeutic countermeasures). In **Figure 2**, re-analyzing the 1P according to the 1P-enriched 3P analyses is illustrated by a double-headed arrow connecting 1P and 3P, signifying the iterative dynamic process connecting the two irreducible phenomena.

Bridge F: Mathematical or Cognitive Modeling

The two irreducible domains have been suggested to be bridged by developing formal mathematical or cognitive models with variables that can refer to either phenomenal or neurophysiological states, an approach previously referred to as "generative passage" (Lutz, 2002). As a general approach to the study of consciousness, this notion has been gaining traction in contemporary theories of consciousness, notably the Integrated Information Theory (Oizumi et al., 2014; Tononi et al., 2016). Other promising theoretical developments build on the free energy principle (Friston, 2009; Friston et al., 2018), offering ways of specifying formal computational models of the autopoietic, embodied and enactive mind (Allen and Friston, 2018; Bruineberg et al., 2018; Kirchhoff, 2018; Ramstead et al., 2018). The concept of predictive processing is here transformed into "predictive engagement" (Gallagher and Allen, 2018), and proposals of how core predictive processing dynamics relate to (pre-reflective) aspects of experience have been put forth (e.g., Seth, 2015; Allen and Tsakiris, 2018; Fabry, 2019; Lutz et al., 2019), calling for rigorous neurophenomenological evaluation. An exceptional project, which puts into practice this approach, is the Projective Consciousness Model (PCM), a mathematical model of embodied consciousness, which is based on the hypothesis that the spatial field of consciousness is structured by a projective geometry and controlled by active inference processes (Rudrauf et al., 2017). While still under development, the PCM helps to account for aspects of subjective character including pre-reflective self consciousness, the 1P point of view, the sense

of ownership, and social self consciousness (Williford et al., 2018), hence providing a mathematical model tying together phenomenological and neural levels of descriptions. In **Figure 2**, Mathematical or cognitive models are illustrated by a double-end inflated arrow connecting 1P and 3P, signifying the need for formal language to connect the two irreducible domains.

Let us close the first part of the paper by suggesting to cease looking for one meaningful bridge between neuronal activity and subjective experience, and rather aim for multiple and diverse feasible bridges. Accordingly, we believe that it is a mistake to think about the NRP as one experiment in which the researcher should choose one of the phenomenological attitudes (thin vs. thick). A fruitful dialogue between 1P and 3P is created by using different kinds of methods on the proposed continuum, at different developmental stages of the NRP, aiming at different insights – each of which can be re-integrated to inform other stages.

One can start a neurophenomenological investigation by implementing the bridge of front-loading (preliminary) phenomenal insights into the NRP study design (bridge a). While executing the study, the phenomenological thickness applied can be guided by various factors such as the number of subjects and available experimental resources, the quality and specificity of the studied phenomenon (in terms of availability and temporal dynamics), the adequacy of available questionnaires and other behavioral measures. Likewise, it is fruitful to use a variety of experimental technologies, as different technologies are useful for different bridges. For example, electrophysiological complexity measures can be suited for comparison with thick phenomenological data and phenomenal validation of neurobiological data (bridge b); specific cognitive tasks might be best suited for measuring underlying mechanisms of specific aspects of experience which can then be analyzed jointly with 1P data (bridge c); neurofeedback is best suited for using physiological data to guide subjective experience (bridge d); and mathematical modeling is highly suitable for creating cognitive models (bridge f). Importantly, bridge e requires a mature NRP, with iterative experiments, and is thus rarely implemented (but is demonstrated in or studies, as detailed subsequently).

In the second part of the paper, we will show how creating a variety of bridges improved our understanding of both the phenomenological side, as well as the neural side, of the embodied self phenomenon we were studying. Yet more importantly, none of these neural or phenomenological aspects by themselves could have led to our current understanding of the embodied, minimal self. The gained insight was a result of re-analyzing 1P data according to the 1P-enriched 3P analyses, representing the maturation of the NRP.

SOME LESSONS FROM OUR JOURNEY WITH NRP: STUDYING SELF-DISSOLUTION

In this section, we lay out a series of studies, demonstrating how harnessing neurophenomenology can advance the study of self consciousness. This direction of inquiry goes back to Varela, Thompson and Rosch's seminal work, The Embodied Mind:

> We believe that mindful awareness practices can provide a natural bridge between cognitive science and human experience (phenomenology). Particularly impressive to us is the convergence that we have discovered among the main themes concerning the self and the relation between subject and object (Varela et al., 1991; The Embodied Mind, p. 33).

We start with a brief review of the field of self consciousness, including basic phenomenological conceptualizations of types of self consciousness, related cognitive and neural counterparts, and methods of scientific inquiry. This is meant to provide a broad context for our series of studies, highlighting their contribution to the readers who are less familiar with this field (while others can skip to the next section).

Studying Self Consciousness

An increasing number of publications in philosophy, psychology and neuroscience investigate "self consciousness" – or the "sense of self," referring here to subjects' consciousness of themselves. The concept of self is highly ambiguous and includes various aspects, thus it may be best construed as a multidimensional construct including somatosensory, agentive, narrative and social components (Gallagher, 2000; Strawson, 2000; Gallagher, 2011, 2013), involving various brain regions (Christoff et al., 2011; Vogeley and Gallagher, 2011; Northoff et al., 2006). As part of a dialogue between philosophy of mind and cognitive neuroscience, a fruitful distinction has been made between two types of processes contributing to the sense of self: self-related and self-specific processing (Legrand and Ruby, 2009; Christoff et al., 2011). The first, self-related processing, attributes or evaluates stimuli with respect to one's perceptual image or mental concept of oneself, giving rise to an enduring sense of identity (such as when contemplating one's own personality, traits, memories or appearance). The second, self-specific processing, specifies the self as an embodied subjective knower and agent. Self-specific features are defined as being exclusive and non-contingent, meaning that they characterize oneself and no-one else, and that changing or losing them entails changing or losing the distinction between 'self' and 'non-self.' Thus, self-specific processes are considered more primal as they implement a functional self/non-self, or self-world distinction in perception, action, cognition, and emotion (Christoff et al., 2011; Seth, 2013). This distinction overlaps with previous differentiations in the literature, such as the 'Me' as opposed to the 'I' (James, 1890), 'extended' vs. 'core' self (Damasio, 1999), and 'narrative' vs. 'minimal' self, respectively (Gallagher, 2000; Gallagher and Zahavi, 2008).

Self-related processes have received the bulk of empirical attention, given that they can be easily manipulated in the lab through cognitive tasks. Neural activations during those tasks overlap strongly with the default-mode network (DMN, Raichle et al., 2001), a large-scale intrinsic network which is highly active at rest (as compared to externally focused goal-directed tasks) as well as during internally focused cognition including self-reflection, episodic memory, future planning, theory of

mind, and personal moral reasoning (Gusnard et al., 2001; Raichle et al., 2001; Northoff et al., 2006; Spreng et al., 2009; Murphy et al., 2018).

Studies of self-specific processing and the minimal self, on the other hand, are less common. Of particular relevance are studies on the neural basis of the senses of agency and ownership (Ionta et al., 2011), the subjective experience of owning and being in control of one's body and thoughts (De Vignemont, 2011; Ionta et al., 2011; Sperduti et al., 2011; Herbert and Pollatos, 2012), as well as the sense of being localized within one's physical body (Blanke and Arzy, 2005; Tsakiris et al., 2008). Cleverly designed experimental setups have managed to create whole body illusions in virtual-reality environments (Alsmith and Longo, 2019), which have been used to directly investigate the experiential and neural implications of manipulating the self-body unity in terms of self-identity, self-location and 1st-person perspective (Kilteni et al., 2012; Serino et al., 2013; Guterstam et al., 2015). Regarding the underlying neuroanatomy, more than any other region, the above studies converge on the right temporoparietal junction and its neighboring regions, involved in multisensory integration and self-other distinction (Donaldson et al., 2015). Another approach to the study of self-specific processing is to investigate real-world cases in which senses of agency or body -ownership appear to be radically disrupted, including psychopathologies such as schizophrenia (Sass, 2013), post-trauma (Ataria, 2018), depersonalization disorder (Gerrans, 2018), and neurological disfunctions involving out-of-body-experience and autoscopy (Blanke and Arzy, 2005).

Importantly, however, an understanding of the neural processes underlying the fully fledged minimal self experience is still lacking. This is due to limitations of the above-mentioned approaches who study local alterations and disruptions of single features of self experience, such as the sense of agency (e.g., thought insertion), body ownership and self-location (e.g., full-body illusions). By contrast, there is emerging empirical evidence suggesting that some non-ordinary states of consciousness involve a more dramatic, global dissolution of the sense of self, and self-specific features in particular (Millière et al., 2018). This might be the case during dreamless sleep (Windt et al., 2016), drug-induced ego dissolution (Letheby and Gerrans, 2017; Millière, 2017) and deep meditative states (as shown by our neurophenomenological studies discussed below: Dor-Ziderman et al., 2013, 2016; Ataria et al., 2015). Of these, the only condition which can be non-chemically and volitionally manipulated in the lab is the deep meditative state. Thus, in addition to meditators' general proficiency in experiential awareness (as discussed above), their specific meditative skills in generating states of global dissolution of self experience render them a uniquely apt cohort for the study.

Our approach to tackling this issue has been as a multidisciplinary team consisting of cognitive neuroscientists, empirical phenomenologists, and in collaboration with an expert meditator (who later became an integral part of the team), who demonstrated in the lab for the first time volitional malleability of the sense of boundaries (SB). Subsequently, we have been exploring meditation-induced neuro-oscillatory and experiential fingerprints of different modes of self consciousness, and 'selfless' states in particular, in highly adept meditators, via the neurophenomenological method (Berkovich-Ohana et al., 2013a; Dor-Ziderman et al., 2013, 2016; Ataria et al., 2015; Berkovich-Ohana, 2015). These studies are subsequently described in some detail.

Previous Neurophenomenological Studies on Self-Dissolution

The studies outlined below demonstrate an evolving research effort, a blueprint, for taking a subject matter which is notoriously difficult to study, and of which very little is known – in terms of both its phenomenology as well as its neural mechanisms – and rendering it tractable to robust scientific investigation. In the present case, this process required three discrete stages to implement a full NRP. The first stage was a proof-of-concept feasibility study in which trained meditators produced deep contemplative states such as timelessness, spacelessness and selflessness under neuroscientific examination (Berkovich-Ohana et al., 2013a; Dor-Ziderman et al., 2013; Berkovich-Ohana, 2015). The second stage (Ataria, 2014; Ataria et al., 2015; Dor-Ziderman et al., 2016) consisted of a zooming-in process in which we: (1) gained more precision on the exact phenomenological construct under study by using thick phenomenological inquiry, (2) developed a suitable experimental setup and research environment, and (3) identified the necessary personnel – both expert meditators and researchers – for carrying out a more refined neurophenomenological study. The third stage, which is a project still underway (see section "A Mature and Comprehensive NRP on SB Dissolution"), lays out a robust, mature and comprehensive neurophenomenological research program centered on sense-of-boundaries dissolution and building on the insights gained from the first two stages. It is important to emphasize that exercising a gradual approach in this project was necessary, given how little was previously known regarding the experiential, as well as neural, dimensions of deep self-dissolution states. And furthermore, given that an established methodology for conducting such studies is still virtually non-existent.

Proof-of-Concept (Building the First Bridges)

The working basis for the study's design was the assumption that long-term Buddhist-oriented mindfulness meditators would be able to produce and hold, volitionally and on demand, certain states pertaining to the self and its dissolution. This assumption was partly based on a preliminary pilot study, which reported two cases of altered states spontaneously occurring during meditation in two proficient practitioners (Berkovich-Ohana, 2015). These states of self dissolution are not uncommon occurrence for insight meditation practitioners and are considered the culmination of mindfulness-induced stages of consciousness. They are characterized by little-to-no conceptual thought and a disintegration of the ordinary subject-object intentional structure of consciousness, which is usually centered on the embodied sense of self. In the Buddhist tradition, these states are deemed highly valuable as they lead to important insights and realizations: "This comprehension of an object noticed, as being impermanent, painful, and

without a self (impersonal) [. . .], by means of simply noticing, without reflecting and reasoning, is called "knowledge by comprehension through direct experience" (Sayadaw, 1964, pp. 10–11; Shulman, 2014)." In this study, participants signaled immediately after the occurrence of such states in the lab, while electroencephalography (EEG) was continuously measured. After the meditation, the participants were asked to freely describe the signaled episodes. The preliminary results demonstrated a unique EEG pattern [an increase in global long-range gamma (25–45 Hz) synchronization] during the signaled states, compared to the background meditation state. Importantly, this preliminary case-study illustrated the feasibility of experiencing spontaneous deep meditative states of self-dissolution in the lab. The phenomenology employed was rather thin, due to the researcher's lack of training in the phenomenological method, yet it allowed the creation of the very first bridge: front-loading 1P insights to 3P study design (bridge a). The neural analysis employed dynamic connectivity within ongoing EEG measurement, thus also enabling phenomenological validation of neurobiological accounts (bridge b).

The next study already recruited a larger cohort of experienced meditators for investigating a range of facets specific to the sense of self. In designing tasks for producing the desired phenomena, we relied on Gallagher's (2000) influential conceptualization of self consciousness as "narrative self" (personal identity with temporal extension) and "minimal self" (momentary awareness rooted in bodily and multisensory processes, endowed with a sense of agency, ownership and 1st person perspective). Our aim was to map the patterns of neural activity underlying narrative and minimal states. However, by front-loading previous phenomenological insights, we also added what we called 'selfless' states, a present-moment awareness devoid of a subjective self core[6]. Rather than define it beforehand, the study's aims were to both characterize this state phenomenologically, as well as to capture its underlying neural fingerprint. The experiment's sample consisted of sixteen long-term meditators tasked with repeatedly producing and holding states pertaining to the narrative self, minimal self, as well as states devoid of the sense of self (Dor-Ziderman et al., 2013). Simultaneously, their brain activity was recorded by magnetoencephalography (MEG), a technology directly measuring the magnetic fields produced by the brain's neurophysiology at a high resolution. It enables differentiating brain activity occurring at different frequency bands including fast brain rhythms, as well as reconstructing their cortical sources. Furthermore, the MEG is setup in a quiet, dark and heavy magnetically shielded room. It is non-invasive and there is no interference from the equipment during the experiment. These factors allow creating a relaxed and intimate environment suitable for the generation of deep meditative states. Each state was produced three times in succession, for 30 s.

As the study was exploratory, and we were not yet experienced with the NRP, we employed retrospective self-reports, as well as two different measures of 1P reports. One was extremely thin, to enable direct naturalization, and the second somewhat thicker:

(a) First, participants evaluated on a 1–3 scale their degree of success in producing each state. The purpose of these numerical reports was to exclude from neural analyses the subjectively non-successful selfless states (ratings of 1). This was done in the MEG immediately after producing each state. Retrospective reports regarding the meditators' perceived (relative to past experiences) success and stability in performing the tasks (on a 1–10 scale, with 1 denoting "very low" and 10 denoting "very high") were collected after the MEG session. Using a similar design, the participants also produced dissolution states in the time and space dimensions (reported in Berkovich-Ohana et al., 2013a). The finding that emerged from these self-report measures was that our participant population, i.e., Buddhist-oriented mindfulness practitioners, were more capable of successfully producing and stably maintaining dissolution states in the self dimension relative to the time and space dimensions (Berkovich-Ohana et al., 2013a; Dor-Ziderman et al., 2013). This finding is coherent with the emphasis placed in Buddhist practice on such experiences. It echoes Varela's suggestion that long-term meditators are especially suitable subjects for studying self experience (Varela et al., 1991), and in particular its subtler pre-reflective aspects (Varela, 2000), within the NRP framework.

(b) Second, following the neural recordings, while still lying within the MEG and via an intercom, participants were asked to briefly describe their experience in the 'selfless' condition freely and in their own words, without reflection or judgment (Jack and Roepstorff, 2002; Schooler, 2002; Lutz and Thompson, 2003). The collected phenomenology was of medium degree on the thin-to-thick continuum due to the lack of skilled phenomenological investigators, as well as technical limitations of conducting interviews during MEG recordings. The short descriptions did, however, make their categorization and validation simpler. The 'selfless' phenomenological descriptions were analyzed and divided into three categories which were then validated by 12 independent judges. These categories indicated either (1) a quieting of experience, (2) an altered dream-like state, and (3) a state lacking sense of ownership or agency. Interestingly, but not surprisingly, the phenomenal categorization was found to be linked with the degree of meditative experience (such that the most experienced meditators were all in the third category). The three phenomenal invariants produced by this process were later used to contrast the third group with the other two, in order to underpin the neural signature of this radical phenomenological shift.

The experiment's neural results indicated that different modes of self-processing involved dissociable frequency-dependent networks (Dor-Ziderman et al., 2013). Narrative, time-extended reflective self-related processing was marked by extensive frontal, and medial prefrontal gamma band (60–80 Hz) power,

[6]The word 'selfless' is used to denote the state of lack of minimal self features, but some authors define it as a minimalist notion of self consciousness which implies its essential reflexivity or first-personal givenness (Zahavi, 2011).

marking attenuation of default-mode activity, in line with fMRI (Whitfield-Gabrieli et al., 2011) and intracranical EEG studies (Nir et al., 2007). In contrast, minimal self processing was linked with beta-band (13–25 Hz) power in a more posterior network including medial (precuneus and posterior cingulate) and inferior parietal lobule (IPL) regions. Contrasting the last phenomenal category with the first two revealed a further right IPL beta-band power reduction, thus linking together phenomenology, meditation experience as well as a distinct neural signature.

To summarize, the neurophenomenological bridging principles we used in this study included: Front-loading 1P insights to the 3P study design – building on the participant's ability to produce 'selfless' states on demand (bridge a); 1P validation of 3P accounts – by collecting phenomenology of the 'selfless' states (bridge b); and joint analyses of 1P and 3P data, by creating three *post hoc* phenomenal categories and using them to contrast the sub-groups and gain new insight which otherwise would not have been available (bridge c).

Zooming-In

The next steps involved thick phenomenology, zooming into the selfless experience with the aim of understanding the phenomenon in terms of both phenomenology and neurophysiology. This thickening of the phenomenological data collection and analysis was enabled by the close collaboration with Y. A., an expert in the phenomenological method.

We began with an in-depth phenomenological study in which 27 advanced mindfulness meditators were interviewed (Ataria, 2014). The goals of this study were:

(1) Mapping the subjective experience during meditation in general terms.
(2) Defining the ability of different meditators to describe their own experience during meditative states (in terms of depth and thickness).
(3) Identifying specific structures in the intentional arc that underwent changes during meditation.
(4) Identifying changes in the meditators' sense of boundaries and sense of self.
(5) Exploring the sense of body ownership during meditation.
(6) Identifying meditators with the ability to volitionally replicate their experience.

At this stage, we were ready for a more detailed case study, examining both the phenomenology and the mediating neural substrate of a well-defined phenomenological construct which emerged from Ataria's (2014) study, namely, the sense of self-world boundaries. At this point, building on insights from an intimate workshop on neurophenomenology[7] which highlighted the necessity to engage in a real and deep dialogue with experts well familiarized with deep contemplative states, we started working in full cooperation with a highly qualified meditator. We were lucky to be connected with a uniquely suited meditator, S.F., a former scientist and well-known meditation teacher with over 40 years and tens of thousands of hours of meditation experience. S.F.'s qualifications were based on both phenomenological as well as neural considerations. Based on phenomenological descriptions provided in previous studies (Berkovich-Ohana et al., 2013a; Dor-Ziderman et al., 2013; Ataria, 2014, 2020) and summarized in Ataria et al. (2015), S.F. stood out as a uniquely apt candidate, able to produce deep meditative states on demand, in a differentiated, replicable and stable manner. In addition, S.F. could describe his experience in clear and precise language, as it was unfolding. Furthermore, S.F.'s neurophysiological data from previous MEG (Berkovich-Ohana et al., 2013a; Dor-Ziderman et al., 2013), EEG (Berkovich-Ohana et al., 2012, 2013b), and fMRI (Berkovich-Ohana et al., 2016a,b) studies on self consciousness, indicated clear and strong effects, which reflected group-level processes (see Ataria et al., 2015). In other words, it was likely that S.F. would be able to produce the required states, describe them in clarity and detail, and that the corresponding neural data would be differentiable between-conditions and generalizable (not idiosyncratic).

The decision to focus on the sense of boundaries (SB) was a result of a discussion between the researchers (Y.A., A. B-O., and S.F.), regarding what would be the best phenomenological dimension that S.F. could alter by demand in the lab. S.F. identified the SB as a phenomenal continuum he could traverse very skillfully, moving along it in a precise way, and stopping in several reliable and differentiated states. SB dissolution is a central goal and skill of Buddhist meditation and has profound implications to the study of self. Albeit a relatively novel research field, there are indications regarding its prevalence among long-term meditators (Lindahl et al., 2017). We decided to focus, for the sake of simplicity, only on 3 highly differentiated phenomenal states ranging from a normal sense of boundaries (SB1) to a state in which the SB was beginning to dissolve (SB2) to a state in which the SB was completely absent (SB3).

The case-study design inherently included the three bridging principles practiced in our previous study: (1) Front-loading 1P insights to 3P study design – building on a series of preliminary phenomenological interviews and discussions with the practitioner S.F., as well as our prior studies; (2) 1P validation of 3P accounts – by collecting phenomenology of all three SB states; and (3) Joint analyses of 1P and 3P data, by creating ad-hoc three phenomenal categories and using them as 'cognitive tasks' to guide the MEG data acquisition and analysis.

The study was set up such that the phenomenological interview was conducted in similar conditions to the subsequent MEG experiment. SF generated the default, dissolving, and disappearing states, SB1, SB2, and SB3, for 1 min each, in succession, for four cycles, while his brain activity was recorded by MEG. We employed thick phenomenology – lengthy in-depth interviews conducted and analyzed by an expert empirical phenomenologist (for more details see Ataria et al., 2015). The thick phenomenology significantly advanced our understanding of the lived experience underlying the 'no-self' state in a number of important respects. While in the previous study care was taken to elicit 1P descriptions of the selfless state, their phenomenology was thin with little detail, richness and specificity. Additionally,

[7] A round table on neurophenomenology (December 2011) with scholars from different fields as well as meditators organized by AB-O and YA at the Inter-Disciplinary Center (IDC), Herzliya, Israel.

the thin methodology was not conducive of the practice of 'bracketing' and thus a layer of Buddhist conceptualization (terms such as 'emptiness,' 'liberation,' and 'witnessing') could still be detected in the descriptions. Third, despite the emphasis placed by Buddhist traditions on 'no-self' experiences as a key to liberation, hardly any phenomenological documentation of such experiences exists due to taboos around discussing such experiences with anyone but one's teacher. The thick analysis revealed SB dissolution experience to be a graded phenomenon, manifesting as nine experiential categories such as diminished or absent sense of agency, ownership, location, egocentric perspective and internal vs. external (Ataria et al., 2015).

In accordance with the diminishing quality of the phenomenal categories, at the level of the brain these changes were mediated only by beta band reductions, with no increases in activity. These beta reductions were localized to bilateral medial and lateral parietal regions (Dor-Ziderman et al., 2016), in particular the right temporal-parietal junction (TPJ, which includes the IPL), in line with our previous study (Dor-Ziderman et al., 2013). These results were coherent with the existing literature as the TPJ, more than any other region, has been shown to mediate the experiential unity of self and body, relying on multisensory integration and contributing to a sense of ownership, agency and self location (Tsakiris et al., 2008; Ionta et al., 2011).

While highly interesting and informative, these results still raised two important concerns. The first regarded the uniqueness of SF in terms of neural patterns. Could these results be generalized to a large population? The second concern regarded the gap between the high complexity of the phenomenology as compared to the neural findings. In other words, while the phenomenology produced nine phenomenal categories, the underlying neural mechanism was linked in previous literature to only some of these categories, so we were unable to discriminate which of the involved SB dissolution phenomenological categories was driving the neural results.

The need to further develop an understanding of the specification of the neural activity related to these experiential changes necessitated a novel study to be designed (the current team project, see section "A Mature and Comprehensive NRP on SB Dissolution"). However, we were already in position to address the first issue based on our previously collected data. By doing so, we implemented for the first time the advanced bridging principle of re-analyzing 1P based on 1P-enriched 3P data (bridge d). Specifically, our better grasp on the SB phenomenology shed new light on group data from our previous study (Dor-Ziderman et al., 2013), and it became apparent that they too demonstrated a form of SB dissolution during the selfless state. Hence, armed with the 1P-enriched 3P data, we could go back to their data and study the exact frequency and regions of interest. As a result, the case study's neural results were partially validated (right hemisphere only) in a larger group (n = 10), and their specificity to the domain of self (not manifesting in control states focusing on the time and space domains) was demonstrated (Dor-Ziderman et al., 2016).

To summarize, this series of studies enabled us to show that it was possible to create, and validate, multiple advanced bridges between thick phenomenology and neuronal activity. The insight gained from both the thick phenomenology, as well as the MEG results, led our team to the third stage, a robust, mature and comprehensive NRP centered on volitional[8] SB dissolution[9]. This project is still underway and is subsequently briefly described.

A Mature and Comprehensive NRP on SB Dissolution

The earlier studies provided phenomenological support for the notion that meditators can profoundly alter their SB in meditation, and the neurophysiological results showed that these alterations were mediated by neural processes linked with embodied self experience in other streams of research (Blanke and Metzinger, 2009). These advancements set the stage for a mature, larger ongoing study aiming to take our NRP one step further. The main aims of this project are (1) specifying underlying neurocognitive models explaining the experiential categories and neural results, and (2) exploring individual differences (neural and phenomenological) borne of the mapping of SB dissolution into phenomenological clusters.

Specifying Underlying Neurocognitive Mechanisms

We aim to specify a cognitive model that is coherent with both the phenomenological and physiological levels of description. Neurophysiologically, this requires the specification of measurable neural parameters relating to key processes within such a model. In the current study, we approach this goal by assessing three candidate neural processes arising from previous empirical research and theoretical reasoning:

(1) The first potential mechanism is the integration of interoceptive signals, previously suggested to give rise to an affectively colored sense of the embodied self (Seth, 2013; Seth and Tsakiris, 2018). This mechanism is indexed in our study using the heartbeat evoked potential, a neurophysiological brain response time-locked to the heartbeat, shown to reflect interoceptive processing of cardiac signals (Schandry and Montoya, 1996; Gray et al., 2007). This measure is recorded and computed during various levels of SB production.

(2) A second potential mechanism is the integration of (motor) efference copies with their actual sensory consequences (re-afferences) (Christoff et al., 2011). Similar and more specific suggestions have been made for the sense of agency (Gallagher, 2000; Haggard and Chambon, 2012), where the suppression of neural responses to self-caused events (as compared to externally caused events) is regarded as the result of efferent/re-afferent integration (i.e., cancelation of sensory changes predicted through efference copies) (Baess et al., 2011). This effect has been shown to correlate with the subjective experience of agency (Gentsch and

[8]It is important to clarify that 'volition' as expressed here is more about a radical letting go than about control; more like a state that can be intentionally cultivated, or recognized.

[9]Volitional SB dissolution cannot be equated with full realization of no-self; rather, it should be viewed as an experiential manifestation of a degree of no-self/emptiness insight.

Schütz-Bosbach, 2011; Timm et al., 2016). This mechanism is indexed in our study using action-induced sensory suppression in a simple task involving button pressing and auditory events, in combination with meditative modulation of the SB.

(3) Finally, a third candidate mechanism is based on multisensory integration accounts of bodily self consciousness (Blanke, 2012; Blanke et al., 2015), which among other methods, has been investigated in peripersonal space paradigms (Blanke et al., 2015; Salomon et al., 2017). We test for peripersonal space modulations during SB dissolution experiences by adapting a previously used neurophysiological multisensory stimulation paradigm (Bernasconi et al., 2018).

By measuring these candidate processes and mapping participants onto related phenomenological dimensions, we aim to establish robust empirical bridges intertwining these two domains.

Individual Differences in Phenomenological Mapping

We aim to explore whether individual differences in phenomenal characteristics map onto different mechanisms. For this purpose, we have recruited a large sample of 50 meditators with a large variance in meditation expertise (115–24,837 h). For creating a mutual phenomenological language regarding the concept of the "sense of boundaries," as well as for increasing the meditators' prospects of successfully producing in the lab clear and stable dissolution experiences, we implemented a 3-week specially tailored meditative training developed and guided by S.F. Following the training, participants were invited to the lab and underwent a varied array of neural and behavioral tasks, phenomenological interviews, self-rating as well as questionnaires. This project is, to our knowledge, the most comprehensive examination to date of the nature of human self-boundaries experience and its neural, behavioral and experiential manifestations.

While the specific training helped focusing participants specifically on this aspect of their meditative practice, the large cohort of meditators entails an unavoidable heterogeneity and richness in participants' meditative experiences of SB alteration. Therefore, an in-depth phenomenological investigation is necessary to make sense of the experiential diversity in a systematic manner. The phenomenological investigation is being conducted using a mixed-methods approach featuring epoch-based self-ratings of stability and depth of the meditative experience (thin phenomenology), follow-up questionnaires and semi-structured qualitative interviews (thick phenomenology). The interviews are conducted based on the open-ended, iterative questioning principles of the Micro-phenomenology method, producing detailed in-depth descriptions of the lived experience of the study participants (Petitmengin, 2006).

Integrating these methods will allow specifying and differentiating the various types of meditative experiences associated with SB alteration, and address the different phenomenological features described by the participants. The interview analysis will capture such diversity by assessing these experiences according to fundamental facets of self-experience such as the sense of location, sense of agency, attentional disposition and affective state. These facets were defined in a top-down fashion, partially based on the previous characterization of SB dissolution (described above), and partially in order to provide subjective parameters corresponding to candidate processes available in the literature (described in the previous section). Within each of these phenomenological categories, there are emerging patterns of variability that characterize and distinguish distinct types of SB dissolution. An additional category that emerged from the preparatory training and preliminary interview analysis is the type of technique (or inner gestures) involved in the dissolution process. Although trained together, participants performed a diverse set of meditative techniques which accordingly produced several distinct variations of the state of SB dissolution. We thus hope that by relying on a larger and diverse sample we can enlighten specifics and commonalities in the enactive dynamics of SB constitution and dissolution. In addition to these thicker aspects of SB dissolution phenomenology, repeated self-ratings throughout the experiments provide temporal tracking of fluctuation in the degree of depth and stability of meditation for each subject.

This full-blown NRP project attempts to implement all the proposed bridges: (1) Front-loading 1P insights to 3P study design – building on the fine-grained phenomenological analysis from the previous case study; (2) 1P validation of 3P accounts – by collecting phenomenology of reduced and enhanced SB states; and (3) Joint analyses of 1P and 3P data, by creating *post hoc* phenomenal categories and using them to guide the MEG data analysis; (4) Using 3P data to define 1P – by using cognitive tasks which engage different aspects of the embodied self (e.g., peripersonal space), we harness the accumulated 3P knowledge related to these tasks to constrain the phenomenology; (5) Using 1P-enriched 3P data to reanalyze 1P – by utilizing the previously found neural markers (in the case study and the small, described in Dor-Ziderman et al., 2016) to refine and build a semi-structured interview focusing on agency, ownership and self location; and (6) Cognitive modeling – we hope to be able to elucidate specific cognitive mechanisms underlying SB flexibility which might, eventually, be integrated into a comprehensive model of embodied self experience.

DISCUSSION

In this paper, we outlined the NRP's requirements, explored its inherent tensions and suggested a typology of bridging principles, constraining the two irreducible domains of 1P and 3P. We then demonstrated the usage of these bridges by describing the unfolding of a series of studies investigating the experience of boundaries of the self, both phenomenologically and neurally. In both realms, the accumulated knowledge was quite limited due to taboos on publicly sharing such 1P accounts and experiences, the difficulty of manifesting such states volitionally, on demand and

under experimental settings, and the lack of suitable cognitive modeling to guide the study. Hence, it is not only that exploring the subtle aspects of self consciousness supports and validates the NRP, but also the reverse, that the NRP is needed to handle such a subtle, profound, fascinating and challenging research topic as conscious experience.

We hope to have been able to demonstrate the potential of the NRP to inform the science of consciousness. We further hope that this account suitably narrates both the challenges and the creative solutions which were needed to be implemented along the way, in order to push this project forward. The guiding intuitions were always in the spirit originally put forth by Varela's NRP of harnessing a pragmatic and flexible stance along the way, collaborating with well-trained participants, and above all, the indispensable need to treat human experience with equal importance as physiological data.

We consider this ongoing circulation between the two realms of physiology and human experience as an act of art, a deep listening, an improvisational dance, which slowly develops into a skillful scientific dialogue. It is not meant for those who use science as a battle to win, or as growing a tail to wag. What is required is passionate teamwork, a willingness to be re-enchanted with the realm of the living and to tackle the mystery of human consciousness from as many angles as possible while practicing pragmatism, flexibility and humility along the way.

AUTHOR CONTRIBUTIONS

AB-O: conceptualization, resources, supervision, investigation, and writing – original draft. YD-Z: methodology, software, data curation, formal analysis, investigation, and writing – original draft. F-MT, YS, and ON: investigation and writing – original draft. SF: writing – review and editing. YA: conceptualization, supervision, investigation, and writing – original draft. All authors contributed to the article and approved the submitted version.

FUNDING

This study was supported by the Bial Foundation grant 45069 to AB-O. F-MT was funded by a DFG individual fellowship (GZ: TR 1587/1-1). YD-Z was funded by an institutional post-doctorate fellowship from the University of Haifa. YS was funded by an excellence Ph.D. fellowship from the University of Haifa. ON was partly funded by Khyentse Foundation.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fpsyg.2020.01680/full#supplementary-material

REFERENCES

Abdoun, O., Zorn, J., Poletti, S., Fucci, E., and Lutz, A. (2019). Training novice practitioners to reliably report their meditation experience using shared phenomenological dimensions. *Conscious. Cogn.* 68, 57–72. doi: 10.1016/j.concog.2019.01.004

Allen, M., and Friston, K. J. (2018). 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., Smallwood, J., Christensen, J., Gramm, D., Rasmussen, B., Gaden Jensen, C., et al. (2013). The balanced mind: the variability of task-unrelated thoughts predicts error-monitoring. *Front. Hum. Neurosci.* 7:743. doi: 10.3389/fnhum.2013.00743

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*, eds M. Tsakiris, and H. D. Preester (Oxford University Press), 27–45.

Alsmith, A. J., and Longo, M. R. (2019). "Using VR technologies to investigate the flexibility of human self-conception," in *About Methodological Advances in Experimental Philosophy* eds Fischer and M. Curtis (London: Bloomsbury). 153–174

Ataria, Y. (2014). Where do we end and where does the world begin? The case of insight meditation. *Philos. Psychol.* 28, 1128–1146. doi: 10.1080/09515089.2014.969801

Ataria, Y. (2017). The answer to the ultimate question of life, the universe, and everything, or: some reflections on the feasibility of the neurophenomenology research program. *J. Conscious. Stud.* 24, 7–30.

Ataria, Y. (2018). *Body Disownership in Complex Posttraumatic Stress Disorder*. Berlin: Springer.

Ataria, Y. (2020). *Consciousness in Flesh*. New York, NY: Palgrave Macmillan.

Ataria, Y., Dor-Ziderman, Y., and Berkovich-Ohana, A. (2015). How does it feel to lack a sense of boundaries? A case study of a long-term mindfulness meditator. *Conscious. Cogn.* 37, 133–147. doi: 10.1016/j.concog.2015.09.002

Ataria, Y., Lahad, M., and Horovitz, O. (2019). Applying the neurophenomenological approach to the study of trauma: theory and practice. *Construct. Found.* 14, 197–214.

Baess, P., Horváth, J., Jacobsen, T., and Schröger, E. (2011). Selective suppression of self-initiated sounds in an auditory stream: an ERP study. *Psychophysiology* 48, 1276–1283. doi: 10.1111/j.1469-8986.2011.01196.x

Bagdasaryan, J., and Le Van Quyen, M. (2013). Experiencing your brain: neurofeedback as a new bridge between neuroscience and phenomenology. *Front. Hum. Neurosci.* 7:680. doi: 10.3389/fnhum.2013.00680

Baquedano, C., and Fabar, C. (2017). Modeling subjects' experience while modeling the experimental design: a mild-neurophenomenology-inspired approach in the piloting phase. *Construct. Found.* 12, 166–180.

Bayne, T. (2004). Closing the gap? Some questions for neurophenomenology. *Phenomenol. Cogn. Sci.* 3, 349–364. doi: 10.1023/B:PHEN.0000048934.34397

Beaton, M. (2013). Phenomenology and embodied action. *Construct. Found.* 8, 298–313.

Berkovich-Ohana, A. (2015). A case study of a meditation-induced altered state: increased overall gamma synchronization. *Phenomenol. Cogn. Sci.* 16, 91–106. doi: 10.1007/s11097-015-9435

Berkovich-Ohana, A. (2017). Radical neurophenomenology: we cannot solve the problems using the same kind of thinking we used when we created them. *Construct. Found.* 12, 156–159.

Berkovich-Ohana, A., Dor-Ziderman, Y., Glicksohn, J., and Goldstein, A. (2013a). Alterations in the sense of time, space and body in the mindfulness-trained brain: a neurophenomenologically-guided MEG study. *Front. Psychol.* 4:912. doi: 10.3389/fpsyg.2013.00912

Berkovich-Ohana, A., Glicksohn, J., and Goldstein, A. (2012). Mindfulness-induced changes in gamma band activity – implications for the default mode network, self-reference and attention. *Clin. Neurophysiol.* 123, 700–710. doi: 10.1016/j.clinph.2011.07.048

Berkovich-Ohana, A., Glicksohn, J., and Goldstein, A. (2013b). Studying the default mode and its Mindfulness-induced changes using EEG functional connectivity. *Soc. Cogn. Affect. Neurosci.* 5, 1–9. doi: 10.1093/scan/nst153

Berkovich-Ohana, A., Harel, M., Hahamy, A., Arieli, A., and Malach, R. (2016a). Alterations in Task-induced activity and resting-state fluctuations in visual and DMN areas revealed in long-term meditators. *NeuroImage* 135, 125–134. doi: 10.1016/j.neuroimage.2016.04.024

Berkovich-Ohana, A., Harel, M., Hahamy, A., Arieli, A., and Malach, R. (2016b). Data for default network reduced functional connectivity in meditators, negatively correlated with meditation expertise. *Data Brief.* 8, 910–914. doi: 10.1016/j.dib.2016.07.015

Bernasconi, F., Noel, J., Park, H. D., Faivre, N., Seeck, M., Spinelli, L., et al. (2018). Audio-tactile and peripersonal space processing around the trunk in human parietal and temporal cortex : an intracranial EEG study. *Cereb. Cortex* 28, 3385–3397. doi: 10.1093/cercor/bhy156

Bielas, J., and Michalczyk, L. (2017). The view from within the brain: does neurofeedback close the gap? *J. Conscious. Stud.* 24, 133–155. doi: 10.18061/dsq.v34i2.4258

Bitbol, M. (2012). Neurophenomenology, an ongoing practice of/in consciousness. *Construct. Found.* 7, 165–173

Bitbol, M. (2019a). Consciousness, being and life: phenomenological approaches to mindfulness. *J. Phenomenol. Psychol.* 50, 127–161. doi: 10.1163/15691624-12341360

Bitbol, M. (2019b). "Neurophenomenology of surprise," in *Surprise at the Intersection of Phenomenology and Linguistics* eds Depraz, N., and Celle, A (Amsterdam: John Benjamins Publishing Company).

Bitbol, M., and Antonova, E. (2016). On the too often overlooked radicality of neurophenomenology. *Construct. Found.* 11, 310–312.

Bitbol, M., and Petitmengin, C. (2016a). "Neurophenomenology and the elicitation interview," in *The Blackwell Companion to Consciousness* 2nd Edn, ed M. Velmans (Hoboken, NJ: Wiley & Sons).

Bitbol, M., and Petitmengin, C. (2016b). On the possibility and reality of introspection. *Mind Matter* 14, 51–75.

Blanke, O. (2012). Multisensory brain mechanisms of bodily self-consciousness. *Nat. Rev. Neurosci.* 13, 556–571. doi: 10.1038/nrn3292

Blanke, O., and Arzy, S. (2005). The out-of-body experience: disturbed self-processing at the temporo-parietal junction. *Neuroscientist* 11, 16–24. doi: 10.1177/1073858404270885

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

Bockelman, P., Reinerman-Jones, L., and Gallagher, S. (2013). Methodological lessons in neurophenomenology: review of a baseline study and recommendations for research approaches. *Front. Hum. Neurosci.* 7:608. doi: 10.3389/fnhum.2013.00608

Bruineberg, J., Kiverstein, J., and Rietveld, E. (2018). The anticipating brain is not a scientist: the free-energy principle from an ecological-enactive perspective. *Synthese* 195, 2417–2444. doi: 10.1007/s11229-016-1239-1

Cardeña, E., Jönsson, P., Terhune, D. B., and Marcusson-Clavertz, D. (2013). The neurophenomenology of neutral hypnosis. *Cortex* 49, 375–385. doi: 10.1016/j.cortex.2012.04.001

Chalmers, D. J. (1995). The puzzle of conscious experience. *Sci. Am.* 273, 80–87.

Chaminade, T., and Decety, J. (2002). Leader or follower? Involvement of the inferior parietal lobule in agency. *Neuroreport* 13, 1975–1978. doi: 10.1097/00001756-200210280-00029

Christoff, K., Cosmelli, D., Legrand, D., and Thompson, E. (2011). Specifying the self for cognitive neuroscience. *Trends Cogn. Sci.* 15, 104–112. doi: 10.1016/j.tics.2011.01.001

Christoff, K., Gordon, A. M., Smallwood, J., Smith, R., and Schooler, J. W. (2009). Experience sampling during fMRI reveals default network and executive system contributions to mind wandering. *Proc. Natl. Acad. Sci. U.S.A.* 106, 8719–8724. doi: 10.1073/pnas.0900234106

Colombetti, G. (2013). Some ideas for the integration of neurophenomenology and affective neuroscience. *Construct. Found.* 8, 288–297.

Damasio, A. R. (1999). *The Feeling of What Happens: Body and Emotion in the Making of Consciousness*. San Diego, CA: Harcourt Brace.

De Haan, S., Rietveld, E., Stokhof, M., and Denys, D. (2013). The phenomenology of deep brain stimulation-induced changes in OCD: an enactive affordance-based model. *Front. Hum. Neurosci.* 7:653. doi: 10.3389/fnhum.2013.00653

De Preester, H. (2002). Naturalizing husserlian phenomenology: an introduction. *Psychoanal. Perspect.* 20, 633–647.

De Vignemont, F. (2011). Embodiment, ownership and disownership. *Conscious. Cogn.* 20, 82–93. doi: 10.1016/j.concog.2010.09.004

Dennett, D. C. (1991). *Consciousness Explained*. Boston, MA: Little, Brown

Depraz, N. (2019). Epoché in light of samatha-vipassana meditation: chögyam trungpa's buddhist teaching facing husserl's phenomenology. *J. Conscious. Stud.* 26, 49–69.

Depraz, N., and Desmidt, T. (2018). Cardiophenomenology: a refinement of neurophenomenology. *Phenomenol. Cogn. Sci.* 18, 493–507. doi: 10.1007/s11097-018-9590-y

Depraz, N., Gyemant, M., and Desmidt, T. (2017). A first-person analysis using third-person data as a generative method a case study of surprise in depression. *Construct. Found.* 12, 190–203.

Depraz, N., Varela, F. J., and Vermersch, P. (2003). *On Becoming Aware: A Pragmatics of Experiencing*. Amsterdam: John Benjamins Publishing.

Desbordes, G., and Negi, L. T. (2013). A new era for mind studies: training investigators in both scientific and contemplative methods of inquiry. *Front. Hum. Neurosci.* 7:741. doi: 10.3389/fnhum.2013.00741

Diaz, B. A., Van Der Sluis, S., Moens, S., Benjamins, J. S., Migliorati, F., Stoffers, D., et al. (2013). The Amsterdam resting-state questionnaire reveals multiple phenotypes of resting-state cognition. *Front. Hum. Neurosci.* 7:446. doi: 10.3389/fnhum.2013.00446

Donaldson, P. H., Rinehart, N. J., and Enticott, P. G. (2015). Noninvasive stimulation of the temporoparietal junction: a systematic review. *Neurosci. Biobehav. Rev.* 55, 547–572. doi: 10.1016/j.neubiorev.2015.05.017

Dor-Ziderman, Y., Berkovich-Ohana, A., Glicksohn, J., and Goldstein, A. (2013). Mindfulness-induced selflessness: a MEG neurophenomenological study. *Front. Hum. Neurosci.* 7:582. doi: 10.3389/fnhum.2013.00582

Dor-Ziderman, Y, Ataria, Y., Fulder, S., Goldstein, A., and Berkovich-Ohana, A. (2016). Self-specific processing in the meditating brain: a MEG neurophenomenological study. *Neurosci. Conscious.* 1, 1–13. doi: 10.1093/nc/niw019

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

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

Fazelpour, S., and Thompson, E. (2015). The Kantian brain: brain dynamics from a neurophenomenological perspective. *Curr. Opin. Neurobiol.* 31, 223–229. doi: 10.1016/j.conb.2014.12.006

Fox, K. C., Nijeboer, S., Solomonova, E., Domhoff, G. W., and Christoff, K. (2013). Dreaming as mind wandering: evidence from functional neuroimaging and first-person content reports. *Front. Hum. Neurosci.* 7: 412. doi: 10.3389/fnhum.2013.00412

Fox, K. C. R., Zakarauskas, P., Dixon, M., Ellamil, M., Thompson, E., and Christoff, K. (2012). Meditation experience predicts introspective accuracy. *PLoS One* 7:e45370. doi: 10.1371/journal.pone.0045370

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. J., Fortier, M., and Friedman, D. A. (2018). Of woodlice and men: a Bayesian account of cognition, life and consciousness. *ALIUS Bull.* 2, 17–43.

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

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. (2011). *The Oxford Handbook of the Self*. Oxford University Press: Oxford.

Gallagher, S. (2012). "On the possibility of naturalizing phenomenology," in Oxford Handbook of Contemporary Phenomenology ed D. Zahavi (Oxford: Oxford University Press) 70-93.

Gallagher, S. (2013). A pattern theory of self. *Front. Hum. Neurosci.* 7:443. doi: 10.3389/fnhum.2013.00443

Gallagher, S., and Allen, M. (2018). Active inference, enactivism and the hermeneutics of social cognition. *Synthese* 195, 2627–2648. doi: 10.1007/s11229-016-1269-8

Gallagher, S., and Sørensen, J. B. (2006). Experimenting with phenomenology. *Conscious. Cogn.* 15, 119–134. doi: 10.1016/j.concog.2005.03.002

Gallagher, S., and Zahavi, D. (2008). *The Phenomenological Mind: An Introduction to Philosophy of Mind and Cognitive Science*. London: Routhledge. 244.

Garrison, K. A., Santoyo, J. F., Davis, J. H., Thornhill, T. A., Kerr, C. E., and Brewer, J. A. (2013). Effortless awareness: using real time neurofeedback to investigate correlates of posterior cingulate cortex activity in meditators' self-report. *Front. Hum. Neurosci.* 7, 1–9. doi: 10.3389/fnhum.2013.00440

Garrison, K. A., Zeffiro, T. A., Scheinost, D., Constable, R. T., and Brewer, J. A. (2015). Meditation leads to reduced default mode network activity beyond an active task. *Cogn. Affect. Behav. Neurosci.* 15, 712–720. doi: 10.3758/s13415-015-0358-3

Gentsch, A., and Schütz-Bosbach, S. (2011). I did it: unconscious expectation of sensory consequences modulates the experience of self-agency and its functional signature. *J. Cogn. Neurosci.* 23, 3817–3828. doi: 10.1162/jocn_a_00012

Gerrans, P. (2018). Depersonalization disorder, affective processing and predictive coding. *Rev. Philos. Psychol.* 10, 401–418. doi: 10.1007/s13164-018-0415-2

Gordon, S. (2013). *Neurophenomenology and its Applications to Psychology*. Berlin: Springer.

Gray, M. A., Taggart, P., Sutton, P. M., Groves, D., Holdright, D. R., Bradbury, D., et al. (2007). A cortical potential reflecting cardiac function. *Proc. Natl. Acad. Sci. U.S.A.* 104, 6818–6823. doi: 10.1073/pnas.0609509104

Guggisberg, A. G., and Mottaz, A. (2013). Timing and awareness of movement decisions: does consciousness really come too late? *Front. Hum. Neurosci.* 7:385. doi: 10.3389/fnhum.2013.00385

Gusnard, D. A., Akbudak, E., Shulman, G. L., and Raichle, M. E. (2001). Medial prefrontal cortex and self-referential mental activity: relation to a default mode of brain function. *Proc. Natl. Acad. Sci. U.S.A.* 98, 4259–4264. doi: 10.1073/pnas.071043098

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

Haggard, P., and Chambon, V. (2012). Sense of agency. *Curr. Biol.* 22, R390–R392. doi: 10.1016/j.cub.2012.02.040

Herbert, B. M., and Pollatos, O. (2012). The body in the mind: on the relationship between interoception and embodiment. *Top. Cogn. Sci.* 4, 692–704. doi: 10.1111/j.1756-8765.2012.01189.x

Hurlburt, R. T., and Akhter, S. A. (2006). The descriptive experience sampling method. *Phenomenol. Cogn. Sci.* 5, 271–301. doi: 10.1007/s11097-006-9024-0

Hurlburt, R. T., Alderson-Day, B., Fernyhough, C., and Kühn, S. (2017). Can inner experience be apprehended in high fidelity? Examining brain activation and experience from multiple perspectives. *Front. Psychol.* 8:43. doi: 10.3389/fpsyg.2017.00043

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

Hurlburt, R. T., and Heavey, C. L. (2015). Investigating pristine inner experience: implications for experience sampling and questionnaires. *Conscious. Cogn.* 31, 148–159. doi: 10.1016/j.concog.2014.11.002

Hurlburt, R. T., Koch, M., and Heavey, C. L. (2002). Descriptive experience sampling demonstrates the connection of thinking to externally observable behavior. *Cogn. Ther. Res.* 26, 117–134.

Husserl, E. (1970). *The Crisis of European Sciences and Transcendental Phenomenology. (D. Carr, Trans.)*. Evanston: Northwestern University Press.

Ionta, S., Heydrich, L., Lenggenhager, B., Mouthon, M., Fornari, E., Chapuis, D., et al. (2011). Multisensory mechanisms in temporo-parietal cortex support self-location and first-person perspective. *Neuron* 70, 363–374. doi: 10.1016/j.neuron.2011.03.009

Jack, A. I., and Roepstorff, A. (2002). Introspection and cognitive brain mapping: from stimulus–response to script–report. *Trends Cogn. Sci.* 6, 333–339. doi: 10.1016/S1364-6613(02)01941-1

James, W. (1890). *The Principles of Psychology*. Vol. 1. Illinois, IL: Dover.

Jo, H-G., Hinterberger, T., Wittmann, M., and Schmidt, S. (2015). Do meditators have higher awareness of their intentions to act? *Cortex* 65, 149–158. doi: 10.1016/j.cortex.2014.12.015

Jo, H-G., Wittmann, M., Borghardt, T. L., Hinterberger, T., and Schmidt, S. (2014). First-person approaches in neuroscience of consciousness: brain dynamics correlate with the intention to act. *Conscious. Cogn.* 26, 105–116. doi: 10.1016/j.concog.2014.03.004

Khachouf, O. T., Poletti, S., and Pagnoni, G. (2013). The embodied transcendental: a Kantian perspective on neurophenomenology. *Front. Hum. Neurosci.* 7:611. doi: 10.3389/fnhum.2013.00611

Kilteni, K., Groten, R., and Slater, M. (2012). The sense of embodiment in virtual reality. *Pres. Teleoperat. Virt. Environ.* 21, 373–387. doi: 10.1162/pres_a_00124

Kirchhoff, M. D. (2018). Autopoiesis, free energy, and the life–mind continuity thesis. *Synthese* 195, 2519–2540. doi: 10.1007/s11229-016-1100-6

Kirchhoff, M. D., and Hutto, D. D. (2016). Never mind the gap. *Construct. Found.* 11, 302–309.

Kordeš, U., and Demšar, E. (2018). Excavating belief about past experience: experiential dynamics of the reflective act. *Construct. Found.* 13, 219–229.

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

Kühn, S., Fernyhough, C., Alderson-Day, B., and Hurlburt, R. T. (2014). Inner experience in the scanner: can high fidelity apprehensions of inner experience be integrated with fMRI? *Front. Psychol.* 5:1393. doi: 10.3389/fpsyg.2014.01393

Le Van Quyen, M., and Petitmengin, C. (2002). Neuronal dynamics and conscious experience: an example of reciprocal causation before epileptic seizures. *Phenomenol. Cogn. Sci.* 1, 169–180.

Legrand, D., and Ruby, P. (2009). What is self-specific? Theoretical investigation and critical review of neuroimaging results. *Psychol. Rev.* 116, 252–282.

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

Lifshitz, M., Cusumano, E. P., and Raz, A. (2013). Hypnosis as neurophenomenology. *Front. Hum. Neurosci.* 7:469. doi: 10.3389/fnhum.2013.00469

Lindahl, J. R., Fisher, N. E., Cooper, D. J., Rosen, R. K., and Britton, W. B. (2017). The varieties of contemplative experience: a mixed-methods study of meditation-related challenges in Western Buddhists. *PLoS One* 12:e0176239. doi: 10.1371/journal.pone.0176239

Lutz, A. (2002). Toward a neurophenomenology as an account of generative passages: a first empirical case study. *Phenomenol. Cogn. Sci.* 1, 133–167.

Lutz, A. (2006). Toward a neurophenomenology as an account of generative passages: a first empirical case study. *Phenomenol. Cogn. Sci.* 1, 133–167.

Lutz, A. (2007). Neurophenomenology and the study of self-consciousness. *Conscious. Cogn.* 16, 765–767.

Lutz, A., Dunne, J. D., and Davidson, R. J. (2007). "Meditation and the neuroscience of consciousness: an introduction," in *The Cambridge Handbook of Consciousness* eds P. D. Zelazo, M. Moscovitch, and E. Thompson (Cambridge: Cambridge University Press) 497–549

Lutz, A., Jha, A., Dunne, J. D., and Saron, C. D. (2015). Investigating the phenomenological matrix of mindfulness-related practices from a neurocognitive perspective. *Am. Psychol.* 70, 632–658. doi: 10.1037/a0039585

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. U.S.A.* 99, 1586–1591. doi: 10.1073/pnas.032658199

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.

Lutz, A., and Thompson, E. (2003). Neurophenomenology integrating subjective experience and brain dynamics in the neuroscience of consciousness. *J. Conscious. Stud.* 10, 31–52.

Mackenzie, M. J., Carlson, L. E., Paskevich, D. M., Ekkekakis, P., Wurz, A. J., Wytsma, K., et al. (2014). Associations between attention, affect and cardiac activity in a single yoga session for female cancer survivors: an enactive neurophenomenology-based approach. *Conscious. Cogn.* 27, 129–146.

McCall, C., Hildebrandt, L. K., Bornemann, B., and Singer, T. (2015). Physiophenomenology in retrospect: memory reliably reflects physiological arousal during a prior threatening experience. *Conscious. Cogn.* 38, 60–70.

Merleau-Ponty, M. (1996). *Phenomenology of Perception*. French: Motilal Banarsidass Publishe.

Millière, R., Carhart-Harris, R., Roseman, L., Trautwein, F. M., and Berkovich-Ohana, A. (2018). Psychedelics, meditation and self-consciousness. *Front. Psychol.* 9:1475. doi: 10.3389/fpsyg.2018.01475

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.2017.00245

Mrazek, M. D., Franklin, M. S., Phillips, D. T., Baird, B., and Schooler, J. W. (2013). Mindfulness training improves working memory capacity and GRE performance while reducing mind wandering. *Psychol. Sci.* 24, 776–781.

Murphy, C., Jefferies, E., Rueschemeyer, S-A., Sormaz, M., Wang, H-T., Margulies, D. S., et al. (2018). Distant from input: evidence of regions within the default mode network supporting perceptually-decoupled and conceptually-guided cognition. *NeuroImage* 171, 393–401. doi: 10.1016/j.neuroimage.2018.01.017

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

Newen, A., De Bruin, L., and Gallagher, S. (Eds). (2018). *The Oxford Handbook of 4E Cognition*. Oxford: Oxford University Press.

Nielsen, T. (2017). Microdream neurophenomenology. *Neurosci. Conscious.* 2017, 1–17. doi: 10.1093/nc/nix001

Nir, Y., Fisch, L., Mukamel, R., Gelbard-Sagiv, H., Arieli, A., Fried, I., et al. (2007). Coupling between neuronal firing rate, gamma LFP, and BOLD fMRI is related to interneuronal correlations. *Curr. Biol.* 17, 1275–1285.

Northoff, G., and Heinzel, A. (2006). First-Person Neuroscience: a new methodological approach for linking mental and neuronal states. *Philos. Ethics Human. Med.* 1, 1–10.

Northoff, G., Heinzel, A., de Greck, M., Bermpohl, F., Dobrowolny, H., and Panksepp, J. (2006). Self-referential processing in our brain—A meta-analysis of imaging studies on the self. *NeuroImage* 31, 440–457.

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

Olivares, F. A., Vargas, E., Fuentes, C., Martínez-Pernía, D., and Canales-Johnson, A. (2015). Neurophenomenology revisited: second-person methods for the study of human consciousness. *Front. Psychol.* 6:673. doi: 10.3389/fpsyg.2015.00673

Ollagnier-Beldame, M., and Cazemajou, A. (2019). Intersubjectivity in first encounters between healthcare practitioners and patients: micro-phenomenology as a way to study lived experience. *Human. Psychol.* 47, 404–425

Overgaard, M. (2004). On the naturalising of phenomenology. *Phenomenol. Cogn. Sci.* 3, 365–379. doi: 10.1023/B:PHEN.0000048939.62282

Parkin, L., Morgan, R., Rosselli, A., Howard, M., Sheppard, A., Evans, D., et al. (2014). Exploring the relationship between mindfulness and cardiac perception. *Mindfulness* 5, 298–313.

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.

Petitmengin, C. (2017). Enaction as a lived experience: towards a radical neurophenomenology. *Construct. Found.* 12, 139–165. doi: 10.1016/j.bbmt.2013.04.012

Petitmengin, C., and Lachaux, J. P. (2013). Microcognitive science: bridging experiential and neuronal microdynamics. *Front. Hum. Neurosci.* 7:617. doi: 10.3389/fnhum.2013.00617

Petitmengin, C., Navarro, V., and Le Van Quyen, M. (2007). Anticipating seizure: pre-reflective experience at the center of neuro-phenomenology. *Conscious. Cogn.* 16, 746–764.

Petitmengin, C., Remillieux, A., and Valenzuela-Moguillansky, C. (2019a). Discovering the structures of lived experience. *Phenomenol. Cogn. Sci.* 18, 691–730.

Petitmengin, C., van Beek, M., Bitbol, M., Nissou, J. M., and Roepstorff, A. (2019b). Studying the experience of meditation through micro-phenomenology. *Curr. Opin. Psychol.* 28, 54–59.

Petitmengin, C., van Beek, M., Bitbol, M., and Nissou, J. M. (2017). What is it like to meditate? Methods and issues for a micro-phenomenological description of meditative experience. *J. Conscious. Stud.* 24, 170–198.

Petitmengin, C., Baulac, M., and Navarro, V. (2006). Seizure anticipation: are neurophenomenological approaches able to detect preictal symptoms? *Epilepsy Behav.* 9, 298–306.

Petitmengin, C. (ed.) (2011). *Ten Years' Viewing From Within: Further Debate*. Thorverton: Imprint Academic.

Petitot, J., Varela, F. J., Pacoud, B., and Roy, J-M. (1999). "Beyond the gap: an introduction to naturalizing phenomenology," in *Naturalizing Phenomenology* eds J. Petitot, F. J. Varela, B. Pacoud, and J.-M. Roy (Stanford, CA: Stanford University Press). 1–80

Polkinghorne, D. E. (1989). "Phenomenological research methods," in *Existential-Phenomenological Perspectives in Psychology* eds Valle R.S., Halling S (Boston, MA: Springer), 41–60.

Price, D. D., and Barrell, J. J. (2012). *Inner Experience and Neuroscience: Merging Both Perspectives*. MIT Press.

Price, D. D., Barrell, J. J., and Rainville, P. (2002). Integrating experiential–phenomenological methods and neuroscience to study neural mechanisms of pain and consciousness. *Conscious. Cogn.* 11, 593–608.

Raichle, M. E., MacLeod, A. M., Snyder, A. Z., Powers, W. J., Gusnard, D. A., and Shulman, G. L. (2001). A default mode of brain function. *Proc. Natl. Acad. Sci. U.S.A.* 98, 676–682.

Rainville, P., Duncan, G. H., Price, D. D., Carrier, B., and Bushnell, M. C. (1997). Pain affect encoded in human anterior cingulate but not somatosensory cortex. *Science* 277, 968–971.

Ramstead, M. J. D., Badcock, P. B., and Friston, K. J. (2018). Answering Schrödinger's question: a free-energy formulation. *Phys. Life Rev.* 24, 1–16. doi: 10.1016/j.plrev.2017.09.001

Reinerman-Jones, L., Sollins, B., Gallagher, S., and Janz, B. (2013). Neurophenomenology: an integrated approach to exploring awe and wonder. *South Afr. J. Philos.* 32, 295–309.

Rudrauf, D., Bennequin, D., Granic, I., Landini, G., Friston, K., and Williford, K. (2017). A mathematical model of embodied consciousness. *J. Theor. Biol.* 428, 106–131. doi: 10.1016/j.jtbi.2017.05.032

Rudrauf, D., Lutz, A., Cosmelli, D., Lachaux, J. P., and Le Van Quyen, M. (2003). From autopoiesis to neurophenomenology: Francisco Varela's exploration of the biophysics of being. *Biol. Res.* 36, 27–66.

Salomon, R., Noel, J. P., Łukowska, M., Faivre, N., Metzinger, T., Serino, A., et al. (2017). Unconscious integration of multisensory bodily inputs in the peripersonal space shapes bodily self-consciousness. *Cognition* 166, 174–183. doi: 10.1016/j.cognition.2017.05.028

Sass, L. A. (2013). Self-disturbance and schizophrenia: structure, specificity, pathogenesis. *Rech. Psychanal.* 2, 119–132.

Sayadaw, M. (1964). *The Progress of Insight: A Treatise on Satipatthana Meditation*. Kandy: Buddhist Publication Society.

Schandry, R., and Montoya, P. (1996). Event-related brain potentials and the processing of cardiac activity. *Biol. Psychol.* 42, 75–85.

Schooler, J. W. (2002). Re-representing consciousness: dissociations between experience and meta-consciousness. *Trends Cogn. Sci.* 6, 339–344.

Searle, J. (1992). *The Rediscovery of the Mind*. Cambridge, MA: MIT Press.

Serino, A., Alsmith, A., Costantini, M., Mandrigin, A., Tajadura-Jimenez, A., and Lopez, C. (2013). Bodily ownership and self-location: components of bodily self-consciousness. *Conscious. Cogn.* 22, 1239–1252. doi: 10.1016/j.concog.2013.08.013

Seth, A. K. (2013). Interoceptive inference, emotion, and the embodied self. *Trends Cogn. Sci.* 17, 565–573.

Seth, A. K. (2015). Presence, objecthood, and the phenomenology of predictive perception. *Cogn. Neurosci.* 6, 111–117. doi: 10.1080/17588928.2015.1026888

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

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

Solomonova, E., Fox, K. C., and Nielsen, T. (2014). Methodological considerations for the neurophenomenology of dreaming: commentary on Windt's "Reporting dream experience." *Front. Hum. Neurosci.* 8:317. doi: 10.3389/fnhum.2014.00317

Sperduti, M., Tallon-Baudry, C., Hugueville, L., and Pouthas, V. (2011). Time is more than a sensory feature: attending to duration triggers specific anticipatory activity. *Cogn. Neurosci.* 2, 11–18.

Spreng, R. N., Mar, R. A., and Kim, A. S. N. (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

Stewart, J., Gapenne, O., and Di Paolo, E. (2010). Enaction: toward a new paradigm for cognitive science. Cambridge, MI: The MIT Press.

Strawson, G. (2000). "The phenomenology and ontology of the self," in *Exploring the Self: Philosophical and Psychopathological Perspectives on Self-experience* ed D. Zahavi (Amsterdam: John Benjamins Publishing) 39–54.

Strle, T. (2013). Why should we study experience more systematically: neurophenomenology and modern cognitive science. *Interdiscipl. Descript. Compl. Syst.* 11, 376–390.

Stuart, S. A. (2013). The union of two nervous systems: neurophenomenology, enkinaesthesia, and the Alexander technique. *Construct. Found.* 8, 314–323.

Thompson, E. (2006). "Neurophenomenology and contemplative experience," in *The Oxford Handbook of Science and Religion* ed P. Clayton (Oxford: Oxford University Press) 226–235.

Thompson, E., Lutz, A., and Cosmelli, D. (2005). "Neurophenomenology: an introduction for neurophilosophers," in *Cognition and the Brain: The Philosophy and Neuroscience Movement* eds A. Brook, K. Akins, & A. K. Brook (Cambridge: Cambridge University Press) 40–97.

Timm, J., Schoenwiesner, M., Schroeger, E., and SanMiguel, I. (2016). Sensory suppression of brain responses to self-generated sounds is observed with and without the perception of agency. *Cortex* 80, 5–20.

Timmermann, C., Roseman, L., Schartner, M., Milliere, R., Williams, L., Erritzoe, D., et al. (2019). Neural correlates of the DMT experience as assessed via multivariate EEG. *Sci. Rep.* 9:16324. doi: 10.1101/706283

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

Treves, I. N., Tello, L. Y., Davidson, R. J., and Goldberg, S. B. (2019). The relationship between mindfulness and objective measures of body awareness: a meta-analysis. *Sci. Rep.* 9, 1–12. doi: 10.1038/s41598-019-53978-6

Tsakiris, M., Costantini, M., and Haggard, P. (2008). The role of the right temporo-parietal junction in maintaining a coherent sense of one's body. *Neuropsychologia* 46, 3014–3018. doi: 10.1016/j.neuropsychologia.2008.06.004

Valenzuela Moguillansky, C., O'Regan, J. K., and Petitmengin, C. (2013). Exploring the subjective experience of the "rubber hand" illusion. *Front. Hum. Neurosci.* 7:659. doi: 10.3389/fnhum.2013.00659

Valenzuela-Moguillansky, C., and Vásquez-Rosati, A. (2019). An analysis procedure for the micro-phenomenological interview. *Construct. Found.* 14, 123–156.

Van den Bussche, E., Vermeiren, A., Desender, K., Gevers, W., Hughes, G., Verguts, T., et al. (2013). Disentangling conscious and unconscious processing: a subjective trial-based assessment approach. *Front. Hum. Neurosci.* 7:769. doi: 10.3389/fnhum.2013.00769

Van Lutterveld, R., and Brewer, J. (2015). Neurofeedback from the posterior cingulate cortex as a mental mirror for meditation. *Biofeedback* 43, 117–120. doi: 10.5298/1081-5937-43.3.05

Varela, F. J. (1996). Neurophenomenology: a methodological remedy for the hard problem. *J. Conscious. Stud.* 3, 330–349.

Varela, F. J. (1999). "The specious present: a neurophenomenology of time consciousness," in *Naturalizing Phenomenology: Issues in Contemporary Phenomenology and Cognitive Science* eds F. Petitot, J. Varela, J. M. Roy, & B. Pachoud (Stanford, CA: Stanford University Press) 266–314.

Varela, F. J. (2000). "The phenomenology of Sunyata," in *Gnosis, Metaphysique, Phénoménologie* eds N. Depraz and J. F. Marquet (Paris: du Cerf).

Varela, F. J., Thompson, E., and Rosch, E. (1991). *The Embodied Mind: Cognitive Science and Human Experience*. Cambridge, MA: MIT Press.

Varela, J., and Shear, J. (1999). First-person methodologies: what, why, how. *J. Conscious. Stud.* 6, 1–14.

Vermersch, P. (2009). Describing the practice of introspection. *J. Conscious Stud.* 16, 20–57.

Vogeley, K., and Gallagher, S. (2011). "Self in the brain," in *The Oxford Handbook of the Self* ed S. Gallagher (Oxford: Oxford University Press), 111–136.

Vörös, S. (2019). Embodying the non-dual. *J. Conscious. Stud.* 26, 70–94.

Vörös, S., Froese, T., and Riegler, A. (2016). Epistemological odyssey: introduction to special issue on the diversity of enactivism and neurophenomenology. *Construct. Found.* 11, 189–203.

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

Williford, K., Bennequin, D., Friston, K., and Rudrauf, D. (2018). The projective consciousness model and phenomenal selfhood. *Front. Psychol.* 9:2571. doi: 10.3389/fpsyg.2018.02571

Windt, J. M., Nielsen, T., and Thompson, E. (2016). Does consciousness disappear in dreamless sleep? *Trends Cogn. Sci.* 20, 871–882

Zahavi, D. (2002). First-person thoughts and embodied self-awareness: some reflections on the relation between recent analytical philosophy and phenomenology. *Phenomenol. Cogn. Sci.* 1, 7–26.

Zahavi, D. (2004). Phenomenology and the project of naturalization. *Phenomenol. Cogn. Sci.* 3, 331–347.

Zahavi, D. (2011). "The experiential self: objections and clarifications," in *Self, No Self* eds M. Siderits, E. Thompson, D. Zahavi (Oxford: Oxford University Press) 56–78.

Zanesco, A. P., King, B., MacLean, K., and Saron, C. D. (2013). Executive control and felt concentrative engagement following intensive meditation training. *Front. Hum. Neurosci.* 7:566. doi: 10.3389/fnhum.2013.00566

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: 24 July 2020
doi: 10.3389/fpsyg.2020.01676

Empirical Evidence for Intraspecific Multiple Realization?

*Francesca Strappini[1], Marialuisa Martelli[2], Cesare Cozzo[3] and Enrico di Pace[2]**

[1] Neurobiology Department, Weizmann Institute of Science, Rehovot, Israel, [2] Department of Psychology, Sapienza University of Rome, Rome, Italy, [3] Department of Philosophy, Sapienza University of Rome, Rome, Italy

Despite the remarkable advances in behavioral and brain sciences over the last decades, the mind-body (brain) problem is still an open debate and one of the most intriguing questions for both cognitive neuroscience and philosophy of mind. Traditional approaches have conceived this problem in terms of a contrast between physicalist monism and Cartesian dualism. However, since the late sixties, the landscape of philosophical views on the problem has become more varied and complex. The *Multiple Realization Thesis* (MRT) claims that mental properties can be (or are) realized, and mental processes can be (or are) implemented by neural correlates of different kinds. Thus, MRT challenges the psychoneural type-identity theory and the corresponding reductionism. Many philosophers have acknowledged the *a priori* plausibility of MRT. However, the existence of empirical evidence in favor of intraspecific, human multiple realizations of mental processes and properties is still controversial. Here, we illustrate some cases that provide empirical evidence in support of MRT. Recently, it has been proposed that foveal agnosic vision, like peripheral vision, can be restored by increasing object parts' spacing (Crutch and Warrington, 2007; Strappini et al., 2017b). Agnosic fovea and normal periphery are both limited by crowding, which impairs object recognition, and provides the signature of visual integration. Here, we define a psychological property of restored object identification, and we cross-reference the data of visually impaired patients with different etiologies. In particular, we compare the data of two stroke patients, two patients with posterior cortical atrophy, six cases of strabismic amblyopia, and one case with restored sight. We also compare these patients with unimpaired subjects tested in the periphery. We show that integration (i.e., restored recognition) seems to describe quite accurately the visual performance in all these cases. Whereas the patients have different etiologies and different neural correlates, the unimpaired subjects have no neural damage. Thus, similarity in the psychological property given the differences in the neural substrate can be interpreted in relation to MRT and provide evidence in its support. Finally, we will frame our contribution within the current debate concerning MRT providing new and compelling empirical evidence.

Keywords: multiple realizability, identity theory, visual integration, natural kinds, antireductionism, crowding, object recognition, functionalism

OPEN ACCESS

Edited by:
Wenjing Cai,
Shanghai Jiao Tong University, China

Reviewed by:
John Bickle,
Mississippi State University, United States
Mark Couch,
Seton Hall University, United States

****Correspondence:***
Enrico di Pace
enrico.dipace@uniroma1.it

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

Received: *03 May 2020*
Accepted: *19 June 2020*
Published: *24 July 2020*

Citation:
Strappini F, Martelli M, Cozzo C and di Pace E (2020) Empirical Evidence for Intraspecific Multiple Realization? Front. Psychol. 11:1676. doi: 10.3389/fpsyg.2020.01676

INTRODUCTION

Understanding what the mind is, its nature, and how it relates to the physical matter, the brain, represents one of the most basic and powerful questions through all human history. Nevertheless, for both science and western philosophy, a definitive answer remained elusive. On the one hand, cognitive neuroscientists have tried to address the problem on an empirical basis by studying the brain mechanisms underlying the cognitive functions, such as visual perception, learning, memory, and so on. On the other hand, philosophers of the mind have approached the problem from a broader point of view and have raised questions on how the mind is related to the existence of the body and how does it fit into the natural world.

The mind–body relation as a "problem" can be traced back to seventeenth century French philosopher René Descartes who asked how the material body, which works according to the physical laws, can interact with the immaterial mind. However, it is widely held that the modern debate over the mind–body problem began only later in the 1950s when physicalism, the view that everything that exists has an ultimate physical nature, became the dominant metaphysical perspective. Inspired by physicalism, the British philosopher and psychologist Place (1956), the Austrian philosopher Feigl (1958), and the Australian Philosopher Jack Smart (1959) proposed a view that became popular as the "identity theory," which claims that mental states are identical with physical states in the following sense: for each type of mental state M there is a (finitely specifiable) type of physical state P such that any individual x is in the state M if, and only if, x is in P. According to this view, the mind–body problem is solved by recognizing that any mental state is a neurophysiological state in the nervous system – rather than having its neural correlates. This view is sometimes called psychoneural "type-identity" theory because it is maintained that the relevant type of mental state is a type of neurophysiological/physical state. Moreover, the term "type" highlights that both the mental and neural states are intended to be a general class of events (types), such as the mental state of feeling pain, rather than specific, spatiotemporally individuated instances ("token") of a certain type, such as the same feeling of pain experienced in different conditions and timing.

Although this new approach was in line with the optimistic mood about the role of the modern science of that time and set some basic and useful constraints for future debates on the mind–body problem, it was eventually short-lived. One of the biggest challenges for the identity theory was the new fundamental change in approaching the problem that arose in the philosophic scenario, called functionalism. In 1967, the American philosopher Hilary Putnam, with the paper "Psychological Predicates" and other works, proposed that a mental state is a functional or computational state. So, the mind–body problem was solved by considering the mind neither as a non-physical thing nor as a physical one, such as a neurophysiological state, but rather in relation to its functionality. Following the naive brain–computer analogy, two computers can compute the same task (function) yet have two different physical states or hardware. So, two nervous systems can perform the same mental task (function) yet having two different neurophysiological states. This argument, known as the *Multiple Realization Thesis* (MRT) implies that mental states can be implemented by different neural correlates. Although MRT originated from the theoretical framework of functionalism; nowadays, it is considered separate and independent. In its original conception, MRT was applied across species; Putnam suggested that one mental state, like feeling pain, is likely realized with non-identical neurophysiological states in different animal species, like reptiles, birds, and mollusks (Putnam, 1967). The argument is more or less the following: suppose that a token m is John's being in pain at time t and that m coincides with the excitation of a neural C-fiber (type P); suppose further that m* is a mollusk's being in pain at time t* and that m** is an extraterrestrial creature's being in pain at time t**: the corresponding physical states of the mollusk and of the extraterrestrial creature are of types P* and P**, different from excitation of a C-fiber (i.e., different from P). The tokens m, m*, and m** are all instances of the state M of being in pain, but they correspond to tokens of different physical types: this is what is meant by saying that the mental type M is multiply realized. The argument can be also formulated in terms of mental properties: x is in a state M if, and only if, x has the property of being in M: so, mental properties are multiply realized. The reader may object that this argument may seem plausible, but it has two shortcomings: it lacks sound empirical support, and it concerns different species (one of which is only imaginary).

With this thesis, Putnam suggested that there is not a constant and invariant identity relation between mental and physical states as the identity theory holds. It also challenged all the reductionist approaches that claim that the physical substrate of the mind is exclusively the nervous system.

A few years later, Fodor (1974) extended Putnam's thesis to intraspecies cases. He argued that mental states can have multiple realizations in the nervous system of different individuals that belong to the same species or even in the same individual across different brain states over time. According to this thesis, for (at least some) types of mental states M, there is not a finitely specifiable physical (neural) state P, such that any individual x is in the state M if, and only if, x is in P, because M is multiply realized. In other words, every single token m of type M is a physical (neural) token p of some type, but there is not a finitely specifiable physical type P of which all p's are tokens: different tokens of M correspond to tokens of different physical types.

Although many philosophers acknowledged the plausibility of the argument, clear empirical evidence in favor of intraspecific, human multiple realization of mental states and properties is still missing. On the one hand, findings from cognitive neuroscience have been used both to support and oppose MRT depending on the grain of analysis used (Bechtel and Mundale, 1999). On the other hand, the way that the question of multiple realizations of mental states and properties has been posed, influenced, at least in part, the kinds of answers that have been proposed (Aizawa and Gillett, 2009a).

In this paper, we intend to remedy these shortcomings by describing empirically tested evidence of human multiple realization of mental states. Several formulations of MRT have been proposed in the literature (e.g., Shapiro, 2004; Polger and

Shapiro, 2016). To avoid misunderstandings, one must choose one precise definition of multiple realization. Here, we adopt the formulation proposed by Aizawa and Gillett (2009a; 2009b).

We will present two partially independent cases of multiple realization of a similar psychological property. In both cases, we will capitalize on a well-studied phenomenon, visual crowding, whereby an object cannot be identified in peripheral vision if surrounded by closely spaced elements. Recognition is restored when objects are separated by a range that describes the size of the integration mechanisms responsible for recognition. Thus, our psychological property is defined as a function of a physical parameter in the input stimulus. Based on this clear-cut definition of the property, we can make predictions on when and how the property is realized.

MULTIPLE REALIZATION OF CROWDED OBJECTS IDENTIFICATION

Multiple Realization Definition

In the last two decades, the concept of realization and multiple realization has become the focus of a more stringent analysis. One of the most influential accounts of multiple realization has been proposed by Aizawa and Gillett (2009a; 2009b), stemming from their "dimensioned" framework for realization relations.

According to their multiple realization definition:

A property G is multiply realized if and only if:

(i) under condition \$, an individual s has an instance of property G in virtue of the powers contributed by instances of properties/relations $F_1 - F_n$ to s, or s 's constituents, but not vice versa;

(ii) under condition \$* (which may or may not be identical to \$), an individual s* (which may or may not be identical to s) has an instance of a property G in virtue of the powers contributed by instances of properties/relations $F^*_1 - F^*_m$ to s* or s* ' s constituents, but not vice versa;

(iii) $F_1 - F_n \neq F^*_1 - F^*_m$ and

(iv) under conditions \$ and \$*, $F_1 - F_n$ and $F^*_1 - F^*_m$ are at the same scientific level of properties" (Aizawa and Gillett, 2009a, p. 188).

In this framework, "a property is individuated by the causal powers it potentially contributes to the individuals in which it is instantiated," and the realizer contributes to the power of the property and not vice versa (Aizawa and Gillett, 2009a).

The underlying idea is the following. The realized property G belongs to a certain scientific level or if you prefer, a certain layer of scientifically investigated reality. In our case, G belongs to the psychological level. The realizers, on the other hand, belong to a different, "lower," scientific level. There can be different levels: microphysical, molecular, cellular, etc. We have a realization of G if properties $F_1 - F_n$ at a lower level L determine G under a condition \$ [this is what clause (i) states]. We have another, different, realization of G if different properties $F^*_1 - F^*_m$ at a level L* determine G under a condition \$* [this is what clause (ii) states]. Obviously, to have multiple realization, $F_1 - F_n$ must be different from $F^*_1 - F^*_m$ [this is what clause (iii) states]. However, (i), (ii), and (iii) would be trivially true if the levels L and L* were different (e.g., the cellular level and the microphysical level). That is why, to have genuine multiple realization the level of realizers must be the same, that is, L = L*, and clause (iv) must be fulfilled.

In the next paragraph, we will briefly discuss visual crowding to define our properties.

Visual Crowding

The entire chain of processing from sensation to object recognition is still partially underspecified. Feature detection, the process of filtering perceptually significant elementary units, like edges, is the first step of visual recognition, and it is considered a well-understood phenomenon (Hubel and Wiesel, 1965; Campbell and Robson, 1968; Graham and Nachmias, 1971). The visual system then binds or integrates the detected features to achieve an object representation that enables recognition. The nature of the feature integration process has long been debated. Recently, several studies have shown that a visual phenomenon called "crowding" could shed light on this feature integration process (Pelli et al., 2004; Pelli and Tillman, 2008; Whitney and Levi, 2011). When integration succeeds, the outcome is a correct object recognition; when it fails, we experience crowding, whereby the objects cannot be correctly identified. In this condition, the features mingle together and produce a jumble that is difficult to recognize. You can experience crowding yourself by looking at **Figure 1**. Crowding occurs when the object to identify is surrounded by nearby objects (like letters in words). Recognition is restored if the objects are spaced far enough apart to exceed the integration region (i.e., the area of the visual field in which features are integrated). Visual crowding is an essential bottleneck for object recognition and visual consciousness (crf. review by Levi, 2008). Crowding has pervasive effects in everyday life because most of the time, the majority of the visual scene is crowded, like words in a text.

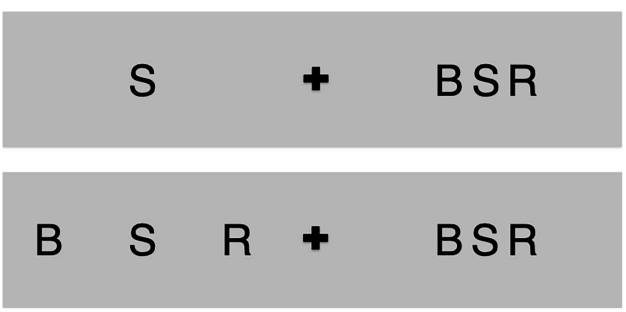


FIGURE 1 | Try to identify the letter S while looking at the central plus in the upper panel. You will see that it is easy on the left and hard on the right. This is crowding: the recognition is hard on the right because the letter S is presented between flankers. Note that this difficulty is not due to acuity, as the targets have the same size. Now look at the plus in the lower panel; again, the recognition is easy on the left and hard on the right. You can escape crowding and restore recognition by increasing the flankers spacing up to a center-to-center distance greater than half of the target viewing eccentricity.

Crowding is considered a mid-level phenomenon that impairs recognition while preserving detection (Pelli et al., 2004). It is operationally defined by psychophysical models that attempt to account for the computation that occurs in the integration region (the region of the visual field where the integration process is computed). These models of crowding suppose feature integration, pooling, source confusion (Treisman and Gelade, 1980; Treisman and Schmidt, 1982; Parkes et al., 2001; Chung et al., 2007; Nandy and Tjan, 2007; Levi, 2008), or a combination of all of these factors (Harrison and Bex, 2017). Pooling refers to compulsory averaging of some elementary feature characteristics, such as orientation, with loss of information about individual elements (Parkes et al., 2001). As for source confusion, this indicates the attribution of one of the object properties to a nearby object, for example, migration of simple oriented elements or color (Treisman and Schmidt, 1982; Nandy and Tjan, 2007).

One essential parameter that characterizes crowding, and that will be important for the definition of our mental property, is the *critical spacing*, the center-to-center spacing between the target object and the flankers needed for recognition (Bouma, 1970). This spacing is proportional to the eccentricity (angular distance from fixation) and independent of object size (Pelli et al., 2004). Specifically, when objects (such as letters or facial features) can be isolated from nearby elements by a critical spacing, features are correctly integrated, and recognition succeeds (Martelli et al., 2005; Chung et al., 2007; Grainger et al., 2010; Rosen et al., 2014; Herzog et al., 2015).

In general, the crowding range, the amount of spacing needed for recognition in any visual field position, defines the size of the integration regions. These regions tile up the entire visual field, but at each retinal location, there is a limit determined by the smallest integration region available at that location. If the smallest available region is too large to isolate the target from the flankers, crowding occurs, and the recognition is impaired. When the smallest integration region available matches the object size excluding the rest, recognition is possible. For this reason, the integration region has also been called *isolation field* to highlight the function of excluding everything that is outside it (Pelli et al., 2004; Martelli et al., 2005). Hereafter, we will use the term crowding and integration interchangeably, as crowding is a ubiquitous, by-product of feature integration.

Some particular types of neuropsychological patients have been reported to require an exceptionally large spacing to restore recognition in foveal vision. Visual crowding was first reported in the foveal vision of strabismic amblyopia and then in normal peripheral vision (Korte, 1923; Irvine, 1945). Recently, it has been shown that also foveal vision in patients with visual agnosia, posterior cortical atrophy (PCA), and visual deprivation is limited by crowding, like peripheral vision in normal subjects (Martelli et al., 2000; Crutch and Warrington, 2009; Strappini et al., 2017b). Surprisingly, these patients have very diverse lesions, all accidental, and sometimes, they do not have any lesion at all. However, they share the same visual behavioral pattern.

In the next paragraphs, we will formalize our evidence of multiple realization by considering two parallel accounts for MRT.

In the first study, we will set the condition requirements according to a stricter definition. The criteria isolate a psychological property characterized by identical input parameters across patients in foveal vision. Specifically, we will compare the performance of some patients with visual impairments in one specific crowding-sensitive task, identification of crowded letters. In all cases, stimuli were presented in the center of the visual field as a function of spacing to show restored recognition. All patients have different etiologies with presumably diverse neural signatures.

In the second study, we will loosen the defining criteria to include both foveal and peripheral vision. In particular, we will present the recognition range restoration in impaired patients and unimpaired subjects tested in their peripheral vision using the same crowding-sensitive task. Although in this case, the condition requirements are more general, the neural substrates are clearly defined as being different based on the retinotopic mapping of foveal and peripheral stimuli.

Study 1: Foveal Crowding in Visually Impaired Patients

Here, we will consider the property of recognizing an object ***o*** placed between two other objects ***a*** and ***b*** **if**, and **only if**:

1. The object ***o*** is presented in foveal vision (same input location across all observers);
2. The objects' size is greater than 0.02 degrees of angle (normal foveal visual acuity);
3. The distance between ***o*** and ***a***, ***o***, and ***b*** needed to restore recognition is greater than 0.07 degrees of angle (the range of normal foveal crowding).

To minimize the effect of crowding, we move the eyes and recognize objects using a retinal foveal region having a diameter of 2 deg (Wandell, 1995). In the first clause, we restrain the analysis only to foveal vision. This region has the highest visual acuity correlated with the smallest scatter and size of the receptive fields of ganglion cells (Hubel and Wiesel, 1974).

In the fovea, acuity (i.e., blur) may impair recognition independently of crowding (Song et al., 2014). Thus, the second clause excludes all the cases in which the limitations on visual recognition depend on visual acuity.

We are constantly surrounded by complex and cluttered scenes, and recognition requires a certain range of critical spacing between two objects (Manassi and Whitney, 2018). Visual crowding has a large critical-spacing range in the periphery and a small one in the fovea, measurable only when the objects are close to acuity (Flom et al., 1963b; Pelli and Rosen, 2015; Coates et al., 2018). So, the third clause is about the critical spacing needed to recognize the target in the fovea when presented in clutter, and it poses an important constraint for defining the cases that will be presented. Specifically, according to the defining criterion, in the patients, we expect recognition to be restored in the fovea with a critical spacing larger than normal, independently from etiology and neural loss (critical spacing >0.07 degrees, Pelli et al., 2016).

Study 2: Foveal Crowding in Visually Impaired Patients and Peripheral Crowding in Unimpaired Subjects

In this second example of multiple realization, we will consider the property of recognizing an object ***o*** placed between two other objects ***a*** and ***b*** **if**:

1. The objects' size is greater than 0.02 degrees of angle (normal foveal visual acuity);
2. The distance between ***o*** and ***a***, ***o***, and ***b*** needed to restore recognition is greater than 0.07 degrees of angle (the range of normal foveal crowding).

This study presents some main differences compared to study 1. First, the deletion of clause 1 from study 1 entails the inclusion of both foveal and peripheral vision. We also turned the biconditional requirement applied in study 1 to a simple conditional statement because clause 2 implies a range greater than 0.07 deg to restore recognition. However, 0.07 deg would be a sufficient range for correct object identification in the normal fovea; thus, the condition in clause 2 is not necessary but only sufficient.

Presentation of the Case Studies

Case 1

LM is a 71-year-old man and retired laboratory technician who suffered an ischemic stroke in the right posterior cortex at the age of 66 years that resulted in left homonymous hemianopsia (visual field loss on the left side of the vertical meridian) (Petersen et al., 2016; Sand et al., 2018).

LB is a retired academic man who suffered a bilateral stroke when he was 81 years old. He has also achromatopsia (color blindness) and topographical disorientation (deficit in navigating familiar external spaces).

Their critical spacing was measured with a recent test developed to study foveal crowding (Pelli et al., 2016). The stimuli were multiple repetitions of a random sample of two letters or digits covering the entire screen. Patients were required to report both while varying inter-item spacing. The spacing threshold (the minimum spacing between stimuli) was measured with an adaptive procedure to reach the 70% accuracy criterion level (see Pelli et al., 2016; Sand et al., 2018 for details). The two patients showed significantly more crowding than a control group (critical spacing $M = 0.175$ degrees of angle [deg], $SD = 0.015$ deg).

Both patients show cortical lesions after the posterior stroke, although with differences in the severity of the extension. For LM, magnetic resonance imaging (MRI) scans acquired 12 months after the stroke showed "an infarction located in the posterior right cerebral hemisphere. Anteriorly, the lesion extends into the right parahippocampal gyrus and, posteriorly, into the lingual gyrus and the medial part of the fusiform gyrus. The lateral portion of the fusiform gyrus is spared, but the white matter above it (the inferior longitudinal fasciculus) is affected. Medially, the lesion surrounds the calcarine sulcus from its most anterior to its most posterior part. There are also two small lacunar infarctions in the right thalamus as well as one in the right centrum semiovale (Sand et al., 2018). LB's MRI scans showed bilateral occipito-temporal infarctions. The right hemisphere lesion extends from the anterior part of the parahippocampal gyrus and might also involve the posterior part of the corpus callosum. Posteriorly, the lesion extends into the lingual and fusiform gyri. In the left hemisphere, there is a small lesion located in the anterior part of the parahippocampal gyrus as well as a small lesion located close to the occipital pole. There are also bilateral white matter lesions in the parietal regions (see Sand et al., 2018 for the MRI images).

Case 2

Patient 1 is a 74-year-old housewife with a diagnosis of PCA, a variant of Alzheimer syndrome characterized by a gradual and progressive deterioration in visual perceptual skills.

Patient 2 is a 58-year-old former care assistant also showing a decline in several perceptual and cognitive tasks compatible with the PCA syndrome (Crutch and Warrington, 2007).

Their crowding range was measured with stimuli composed of target letters flanked by two letters at four spacing conditions (condensed, normal, two-space-expanded, and four-space-expanded). Both patients were severely impaired in this task compared with control subjects (critical spacing: patient 1, 1 deg; patient 2, 1.8 deg).

In patient 1, mild non-specific changes with preserved alpha rhythm were observed with an electroencephalography exam. The MRI scans showed a "mild generalized cerebral atrophy with slightly greater prominence of parietal convexity sulcal spaces." Patient 2 showed an absence of alpha rhythm and extra slow activity in the right temporal region. Visual evoked potentials were normal. The MRI scans showed "mild sulcal widening around the calcarine fissure" (Crutch and Warrington, 2007).

Case 3

Six patients (age: $M = 26.5$, $SD = 14.15$) with amblyopia caused by an early onset of strabism were included in the study (Song et al., 2014). Amblyopia is a condition characterized by a decreased vision in an eye, the input of which is impairly processed by the brain, which over time favors the other eye.

The threshold spacing was measured in the amblyopic eye by varying the inter-letter spacing between a target letter and four flanked letters (above, right, below, and left) with an adaptive procedure to achieve the 50% of accuracy criterion level. All patients were more or less impaired in the crowding test ($M = 1.21$ deg; $SD = 1.44$ deg).

The neural substrates of these patients were not investigated in the referenced study (see section "Discussion" for more details about amblyopia and neural loss).

Case 4

BB was 72 years old at the time of testing. He turned bilaterally blind at the age of 10 years after a violent lime spill that burnt the anterior chamber of both eyes and the posterior of the left eye. Eyes have been sutured to prevent infections. He then studied in Braille and worked as a switchboard operator. At the age of 62 years, after 51 years of complete visual deprivation, he underwent an osteo-odonto-keratoprosthesis intervention to the right eye performed by Prof. Falcinelli at S. Camillo Hospital in

Rome, Italy (Falcinelli, 1993). He recovered sight within a central visual field of 10 degrees and a normal foveal visual acuity of −0.04 logMAR corresponding to 0.07 degrees of visual angle tested with the Snellen eye chart. The visual field restriction is the consequence of the implant's optical characteristics, and it is not due to retinal loss (Falcinelli, 1993). BB visual abilities have been extensively tested after 10 years of recovery (Martelli et al., 2000). He showed normal contrast sensitivity for static and moving gratings with a modest decay for all the spatial frequencies tested compatible with BB's age and impairment in recognizing pictures of objects presented in an unusual perspective (Martelli et al., 2000). BB critical spacing was evaluated centrally measuring contrast threshold for a target letter as a function of the flankers' spacing with an adaptive procedure converging at the 82% accuracy criterion. The critical spacing threshold is identified as the spacing at which letter recognition ability with the flankers equals the ability tested without flankers (i.e., the breakpoint of the function; Pelli et al., 2004). BB shows a large range of foveal crowding whereby recognition is restored if flankers are at a center-to-center spacing of 2.5 deg (Martelli, 2001). The neural substrate of the patient has not been investigated, and BB had no neurological history.

Cases 5

Three non-neurological university students with normal or corrected-to-normal acuity participated in this study (Pelli et al., 2004). Stimuli were presented binocularly at 4 deg of eccentricity, target and flankers size measured 0.32 deg. Apart from this difference, stimuli and procedures were identical to the one used in case 4. Data show that subjects required about 1.2 deg of inter-letter spacing (critical spacing) to restore recognition at the tested eccentricity ($M = 1.28$, $SD = 0.19$).

Comments to Study 1: Foveal Crowding in Visually Impaired Patients

We presented four case studies of patients with visual impairments as explained by foveal crowding. Their visual impairment cannot be explained by early sensory deficits (e.g., low visual acuity or contrast sensitivity), oculomotor disturbances, attentional deficits, aphasic syndromes, and semantic dementia. Thus, referring to clause two and three, we took into account only those subjects whose visual impairment in object recognition might be explained by visual crowding and not acuity (for more details on the relation between visual crowding and acuity, see Song et al., 2014; Strappini et al., 2017b). These patients cannot correctly identify letters if presented foveally in a clutter. To investigate their crowding range, all patients were tested with comparable crowding-sensitive tasks that required the identification of a letter flanked by other letters. All patients obtained a critical spacing greater than 0.07 deg. Thus, patients show a dependency on spacing largely greater than the normal fovea to restore recognition. This sensitivity is a marker of the operation of integrating the visual elements necessary to restore the property of correctly identifying the object. The data show that this operation is realized in the same way in all the reported cases over and above the differences in the patients' etiologies and neuroanatomical impairments.

Comments to Study 2: Foveal Crowding in Visually Impaired Patients and Peripheral Crowding in Unimpaired Subjects

Overall, all the cases we presented are limited by crowding. Both foveal vision in visually impaired patients and peripheral vision in normal subjects require a critical spacing bigger than 0.07 deg to release from crowding and recover recognition. Although all patients presumably have a certain degree of neural impairment, the normal subjects are neurologically intact.

Study 2 identifies a precise correspondence between the critical spacing necessary to restore recognition for the impaired patients tested in foveal vision and the non-neurological observers in peripheral vision at an eccentricity in which the two estimates are equivalent. This comparison is crucial for our thesis because it is known that foveal and peripheral regions from the eyes project to different areas of the brain (see section "Discussion"). Thus, this result leads us to conclude that the psychological property of identifying an object is multiply realized by the foveal pathway in the visual brain of the impaired subjects and the visual peripheral pathway of the non-neurological subjects.

DISCUSSION

Despite several decades of research, the existence of empirical evidence in favor of intraspecific, human, multiple realization of mental processes and properties is still debated. In this study, we consider the psychological property of recognizing an object when presented in clutter a ubiquitous phenomenon of everyday life. In particular, we present two partially different evidence of multiple realization based on some cases of visually impaired patients and normal peripheral vision, both constrained by visual crowding. We show that despite all cases sharing an abnormal range of critical spacing to restore recognition, they highly differ in their neural systems.

In the next paragraphs, we will discuss whether these cases are all possible different realizations of the same psychological property. To that aim, we will first discuss the neural substrates of visual crowding in normal subjects and in the visually impaired patients.

It is crucial for empirical evidence of multiple realization that the neural realizers are at the same neural grain size level (Aizawa and Gillett, 2009a,b). The hierarchy of the nervous system is composed of many levels such as biological macromolecules, synapses, neurons, neural circuits, cortical areas, and systems of areas (e.g., visual system) (Liang et al., 2016, p. 14). In these studies, we will discuss the neural substrates at the level of cortical areas and networks. Here, we informally apply the term *network* to a set of areas that contribute to a particular set of tasks or functions without an explicit reference to the anatomical connections (Petersen and Sporns, 2015). This level has become the most common framework to describe the human cognitive architecture in the last decades (Raichle et al., 2001; Behrmann and Plaut, 2013; Petersen and Sporns, 2015).

Neural Correlates of Crowding in Normal Subjects

It has been suggested that the site of visual integration involves several visual areas in the striate and extrastriate cortex. Most of the research about the neural substrates of the integration process comes from studies of visual crowding on normal subjects tested in peripheral vision. These studies generally agree on locating this phenomenon in the visual cortex beyond the site of binocular combination, based on the observation that there is visual crowding even when the target and the flankers are presented in dichoptic vision (Flom et al., 1963a). However, the precise locus or network is still debated. Some psychophysical and neuroimaging studies have suggested that V1 can be the earliest area showing neural activity modulated by crowding (Anderson et al., 2012; Millin et al., 2013; Chen et al., 2014; Kwon et al., 2014). However, crowding related activation in V1 is absent when attention is diverted away from the stimulus, indicating that V1 involvement in crowding may be the result of feedback suppression coming from higher-order areas (Strappini et al., 2017a).

Evidence for double dissociation of crowding and acuity (Song et al., 2014) suggests that acuity and crowding may be linked to different areas. Acuity is tightly linked to V1, so crowding may be tightly linked to a higher cortical network of regions (Song et al., 2014; Strappini et al., 2017b). Some studies have speculated that V2 has the critical receptive field size to induce crowding (Freeman and Simoncelli, 2011) and that its receptive fields, in synergy with spatial attention, modulate their size to reduce crowding (He et al., 2019). Others have pointed to V3 (Tyler and Likova, 2007; Bi et al., 2009), V4 (Motter, 2006), or higher visual areas (Chung et al., 2007; Freeman et al., 2011). Although the existence of a "crowding area" is still debated, the modulation of the neural activity from early to higher visual areas, like the visual word form area (VWFA), is consistent with the increase in the receptive field size (Freeman et al., 2011; Pelli and Rosen, 2015; Strappini et al., 2017a) and with the occurrence of visual crowding at multiple levels in the visual hierarchy (Whitney and Levi, 2011).

Neural Correlates of Crowding in Visually Impaired Patients

In studying a psychological function, it is possible to obtain useful insights on its characteristics from the study of cases in which that function is impaired. If a neural structure plays a role in the realization of that function, then damage to that neural structure would lead to an impairment of the function.

The cases presented here show some examples of neural implementation of visual crowding. The first two cases, LM and LB (Petersen et al., 2016; Sand et al., 2018) suffered a stroke, a cerebral lesion in which the neuronal death depends on a sudden lack of adequate amount of blood flow, thus oxygen and glucose. Both patients show cortical lesions in the posterior part of the brain, including the visual cortex, specifically in the lingual and fusiform gyri (occipital and temporal lobe, respectively).

The second cases, conversely, have a visual impairment with a gradual onset attributed to PCA syndrome, a progressive neuronal loss in the posterior part of the cerebral cortex. Neuroimaging studies have shown that PCA correlates with severe hypoperfusion in the lateral and medial parieto-occipito-temporal cortices (Kas et al., 2011; Crutch et al., 2017).

Finally, the patients with strabismic amblyopia and restored sight can both be considered as clinical cases of visual deprivation that silence the retinal input from the eye. This is due to suppressive mechanisms on the one end and lack of sensory information on the other. Regarding amblyopia, it is still unknown whether the impairment is due to a feed-forward dominance or feed-back selection of the fellow eye through top–down mechanisms that originate in the extrastriate cortex (Kiorpes and Daw, 2018). Nevertheless, dysfunction in V1 does not seem to be sufficient to explain the visual impairment in amblyopia (Kiorpes et al., 1998; Shooner et al., 2015), whereas the neural correlates of late-blind patients are still unknown.

Overall, the patients are neurologically different in that two show a large loss of the ventral cortex (Sand et al., 2018), PCA patients show general hypoperfusion of the posterior cortex (Crutch et al., 2017), and amblyopes and the visually deprived patient have no evident neuronal loss. In general, it is possible to speculate that in all these patients, visual crowding might reflect a limitation in the number of neurons devoted to foveal integration. Although a few studies have correlated retinal ganglion cell density (Kwon and Liu, 2019) and the number of cortical neurons (Strappini et al., 2017b) to the critical spacing, these clinical cases also point to differences in the way the neuronal decrease can lead to visual crowding.

Foveal Crowding in Visually Impaired Patients as an Instance of Multiple Realization

We can speculate on how visual integration is realized in the visual system as evidence of multiple realization. We should consider whether the neural substrate responsible for recognition at large spacing is the same or different in all these patients. Due to the heterogeneity across all these cases, it is unlikely that an identical neural correlate supports the function. However, we cannot exclude that the recognition ability is supported using the residual functions of the same network.

The network hypothesis requires further consideration about the way in which a network may contribute to the realization of a psychological property in a multiple realization perspective. We will first establish that all the nodes in the network are probably necessary but not sufficient for the network working. If we consider these regions working as "critical" hubs (thus, damage in one area does damage to all the circuit), the neural substrates of the patients would eventually be the same and considered as an example of "merging" of realizers. Several studies have shown that many visual areas are engaged in the representation of multiple functions (cf. review of Behrmann and Plaut, 2013; Kay and Yeatman, 2017). A malfunction of all the network would cause a variety of visual deficits, beyond integration, that has not been observed in the literature (Strappini et al., 2017b). Thus, we can conclude that each node is necessary but not sufficient for the network working.

We have strong physiological evidence that all the visual areas perform different types of processing of the visual input, beyond visual integration, such as detection of elementary features, color and motion perception, and shape processing. Because these nodes are necessary but not sufficient for the network functions, we may speculate on how a partial compromise of those nodes might possibly affect visual integration while sparing the other functions of the nodes.

Just for the sake of speculation, consider the following hypothetical example: it is reasonable to conjecture that an area that detects motion needs integration fields to compute the delay among different events that occur inside those regions. The same computation could be "adapted" to the integration of static visual features, simply setting the interval between the events to zero. This example would again be an instance of multiple realization in that the different areas of a network contribute to computing different features to the realization of the same psychological property, integration. This perspective would be compatible with an even more stringent version of multiple realization that supposes that the realizers are "differently the same" (Shapiro, 2004; Polger and Shapiro, 2016), that is, the differences among the realizers must be relevant to the way the property is realized.

Foveal Crowding in Visually Impaired Patients and Peripheral Crowding in Unimpaired Subjects as an Instance of Multiple Realization

The second study presents a broad comparison between foveal and peripheral crowding as evidence of MRT. The normal range of crowding scales with eccentricity. However, psychophysical models have clearly shown that the computation, integration, is the same across the visual field. This view of visual crowding as an indivisible and homogeneous phenomenon across the visual field contrasts with the high diversity of its neuroanatomical substrate across the foveal and peripheral vision. These neuroanatomical differences are remarkably relevant to MRT. Eccentric retinal regions project to the corresponding cortical areas that represent the peripheral parts of the visual field. This spatial specificity of connections between neurons contributes to the emergence of topographical cortical representations of the visual field (retinotopic maps). In the human primary visual cortex, as an object moves from foveal to peripheral locations of the visual field, the neurons that are activated varies from posterior to anterior parts of the calcarine fissure (Daniel and Whitteridge, 1961; Adams et al., 2007). This organization is preserved in the rest of the retinotopic visual areas (Wandell and Winawer, 2011). Although foveal and peripheral crowding may be associated with the same psychophysical mechanism, this topographical specificity hinders the hypothesis that they are based on the same anatomical structures.

Equating for the crowding range, here, we have shown that recognition is restored similarly in patients' foveal vision as in non-neurological subjects tested in the periphery. Knowing that the neural structures recruited by foveal and peripheral stimuli are different, this is strong evidence for MRT.

Contribution to the Debate on the Empirical Evidence of MR of Psychological Properties

In a paper on the multiple realization of psychological properties, Aizawa and Gillett (2009b) expressed the hope that more scholars would focus their attention on the multiple realization evidence coming from science. Indeed, they are firmly convinced that the discussion on MRT may turn from a traditional theoretical dimension, typical of philosophical debates, to a more concrete empirical evidence-based dimension. From this perspective, we hope that our proposal will contribute to the debate. Next, we will compare our contribution to three attempts to empirically test MRT reported in the literature.

Color Vision

Aizawa and Gillett (2009a) proposed an example of multiple realization based on normal color vision. Chromatic perception depends on the sensitivity to three different primary lights that are processed by three distinct retinal photoreceptors, the short (blue), medium (green), and long (red) wavelength cones. Their different spectral sensitivity is the result of the differences in the chromophore pigments, called opsins, that are contained in these cells. The authors noted that several studies have shown the existence of polymorphisms in the green and red opsins in the normal population. These small variations in the amino acid chains result in slightly different absorption spectra of the opsins, in particular for those codifying red and green. However, these slight variations are included in what is considered normal chromatic perception. Thus, the normal color vision can be considered multiple realized because there are normal variations in its parts.

In this example, the property levels are described in fine detail: the opsin properties multiply realize the photoreceptor properties related to the spectral absorption rate, which is relevant for the property of chromatic perception. However, it has been objected that these polymorphisms may be accounted for normal individual differences (Polger and Shapiro, 2016).

Somehow in line with Polger and Shapiro (2016), we think that in the example provided by Aizawa and Gillet, the cognitive property is missing. We do not exclude *a priori* that color perception (or being trichromat) can be considered a psychological property; however, we think that its phenomenology, its behavioral outcome, is missing from the proposal. We further conjecture that this example could provide concrete evidence of multiple realization if the psychological level was added by showing that there are no differences in color perception among trichromats that have those polymorphisms. Indeed, even slight differences among these normal trichromats would exclude that color vision is multiply realized.

In any case, even if the psychological property would be exactly the same across individuals, the reported evidence of multiple realization would concern low-level peripheral processing (proteins and biological macromolecules). The level at which MRT is usually considered as an alternative to the psychoneural identity theory is the neural level, including areas and networks of areas, which is the focus of our contribution.

Dendritic Spines

Aizawa and Gillett (2009b) discussed in detail another example of multiple realization of a psychological property, this time at the neuronal level. They start from the assumption that probably any psychological property depends in some way on the electrical activity of excitatory neurons across distinct regions of the cortex. This electrical activity may, in turn, be modulated and eventually multiply realized by the properties and relations of other parts of the neural structure. Neurons are notoriously composed of cell bodies, axons, and dendrites; through the activity of synaptic connections, the dendrites receive information from other neurons. The authors focused their attention on a particular substructure of the dendrites, known as dendritic spines. It is believed that dendritic spines may play a role in the memory storage, in the modulation of synaptic strength, and in the transmission of the electrical signal. Dendritic spines have several properties, such as size, length, and volume. It has been shown that such properties may vary over time, from hours to weeks. This neuronal property, that is, transmitting electrical signals to other neurons, seems to be multiply realized by the properties of the dendritic spines – as they vary along a time dimension. Consequently, a psychological property is multiply realized by the properties of the dendritic spines. The authors suggested "remembering something" as an example of a psychological property. Remembering something may remain constant in the same individual, whereas the properties of the dendritic spines vary in time as the properties at the neuronal level.

Although the description of the realizers is very detailed, as noted by the authors, we argue that it cannot be excluded that one day, it will be discovered that the plasticity of the dendritic spines actually does not play any relevant role in the realization of the psychological property, that is, remembering something.

The evidence we provide is not susceptible to the same objection in that the relationship between the psychological property we have described and the associated neural realizers is clearly defined.

Psychopathology

Finally, it is worth mentioning a seminal work recently published by Borsboom et al. (2019). Although the authors' attention is focused on clinical psychology and what may be classified as pathological thoughts, believes, and behaviors, their framework can be applied to mental states in general as well. In the authors' view, the reference to MRT is actually far beyond the limit of the brain and the way in which the neural states realize the psychological properties. Indeed, it extends to a complex network of interconnections between the subject's intentional states (thoughts, desires, and beliefs), the neural states, and the surrounding environment. Briefly, an individual, particular, mental state (e.g., fear) is determined by a coherent pattern of interconnections between the subject's behavior and the surrounding environment (e.g., the subject tries to hide). Generally, we tend to interpret such a pattern of interconnections by making reference to the subjects' intentional contents (e.g., the subject believes that hiding will reduce the fear). Yet, this inference is not enough to understand the exact mental state of the subject and if the behavior is appropriate or dysfunctional. Indeed, the appropriateness of the subject's behavior depends also on cultural and social factors. Consequently, according to the authors, mental states may be realized in many different ways in different people.

Although this framework is conceptually plausible and intriguing, it does not seem to provide compelling evidence in its favor. Let us consider how the desire of taking an umbrella may be realized at the brain level. Under the framework perspective, the desire of taking the umbrella cannot be isolated from a more complex network of related mental contents (e.g., the belief that it will probably rain; the belief that the umbrella will protect you from the rain or sun; the belief that it is not good to expose yourself to the rain or the sun, etc.). These contents may be extremely diverse and idiosyncratic. Thus, it is obviously very unlikely that, in our example, the desire of taking the umbrella would correspond to the same pattern of interconnected intentional states in different subjects, as for the neural realizers of such intentional states. Therefore, the theory simply states that it is highly improbable that the mental contents are realized by the same neural substrates in different subjects.

Here, the theory may be subjected to what has been termed the "Grain-Argument" objection (Bechtel and Mundale, 1999; Aizawa and Gillett, 2009b). According to this argument, although the "grain" at which a psychological property is usually described is coarse, the level at which the supposed neural substrates are described is much finer. As a consequence, a property described vaguely may actually be related to a variety of brain states individuated at a much finer grain. This reasoning may give rise to the illusory impression that the mental property may actually be realized by many different neural substrates. In line with this reasoning, it has been objected that in this framework, the mental properties can be subjected to kind splitting (Pernu, 2017). Reducing the grain at the psychological level may reduce the variability observed at the neuronal level and increase the possible correspondence between the psychological properties and neural realizers.

Although we consider this proposal as a cornerstone that will inspire future promising research attempts at testing MRT at an empirical level, comparing their network theory to our proposal, we may highlight two main differences: (1) our evidence refers more strictly to the relation between the psychological property and the neural realizers; (2) compared to the intentional states, whose nature and status is undoubtedly more complex, we refer to a simpler and indivisible psychological property (i.e., the ability to recognize an object as a function of a physical stimulus parameter spacing), which may turn out to be relevant in response to the Grain-Argument objection and taxonomic "kind splitting" (Polger and Shapiro, 2016).

Is Our Proposal Question-Begging?

Various formulations/definitions of MR may be taken into account to verify whether a property may be multiply realized. Each definition establishes some criteria that should be fulfilled in order for some evidence to be considered a case of MR. To test whether our empirical evidence could be considered an example of MR of a psychological property, we chose the definition of MR proposed by Aizawa and Gillett (2009a). Compared with other

formulations (e.g., Polger and Shapiro, 2016), it is not the most conservative. The comparison among different formulations of MR is surely an interesting issue to be discussed, but it is more theoretical/philosophical in itself and, as a consequence, beyond the limit of the present paper, which is mainly empirical. Under this respect, the fact that the Aizawa and Gillett definition of MR is less conservative does not make it trivial the quest for properties that could be multiply realized. Even if it would turn out that, in the end, everything is multiply realized (which is of course far from obvious), it would remain in any case a question of empirical evidence.

Whether or not it does exist a human intraspecific case that fulfills the conditions proposed by Aizawa and Gillett is an empirical issue, and our paper is in fact aimed at looking for an answer to such an empirical question.

CONCLUSION

In this study, we considered the visual phenomenon of crowding in visually impaired patients and normal subjects as possible evidence of human multiple realization of mental properties. We further discussed the virtues and limits of our proposal compared with some previous empirical evidence reported in the literature.

Although we acknowledge that our evidence is far from conclusive, we think that it provides a fruitful bridge between philosophical and scientific approaches in the study of the relationship between mental properties and the human brain. In particular, we anticipate that our proposal, integrating findings from neuropsychology and psychophysics, will help brain scientists to search for hypothetical multiple realizers by considering the compatibility of their data with the multiple realization view. Consistently, it has been suggested that scientists might already have produced such data, although rarely does the term "realization" appear in their works (Aizawa and Gillett, 2009b). This type of data could be for example the presence of some relevant "outliers," that rather than being a nuisance to regress out might indicate the existence of greater, unknown complexity in the studied phenomenon.

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.

AUTHOR CONTRIBUTIONS

EP had the idea to compare visual crowding and agnosia as a possible example of multiple realization. All the authors contributed to the conjecture and writing.

FUNDING

This work was supported by the Italian Department of Health and the Sapienza University.

ACKNOWLEDGMENTS

We thank Amihai Ben-Nun for useful comments.

REFERENCES

Adams, D. L., Sincich, L. C., and Horton, J. C. (2007). Complete pattern of ocular dominance columns in human primary visual cortex. *J. Neurosci.* 27, 10391–10403. doi: 10.1523/jneurosci.2923-07.2007

Aizawa, K., and Gillett, C. (2009a). "Levels, individual variation, and massive multiple realization in neurobiology," in *The Oxford Handbook of Philosophy and Neuroscience*, ed. J. Bickle (Oxford: Oxford University Press).

Aizawa, K., and Gillett, C. (2009b). The (multiple) realization of psychological and other properties in the sciences. *Mind Lang.* 24, 181–208. doi: 10.1111/j.1468-0017.2008.01359.x

Anderson, E. J., Dakin, S. C., Schwarzkopf, D. S., Rees, G., and Greenwood, J. A. (2012). Te neural correlates of crowding-induced changes in appearance. *Curr. Biol.* 22, 1199–1206. doi: 10.1016/j.cub.2012.04.063

Bechtel, W., and Mundale, J. (1999). Multiple realizability revisited: linking cognitive and neural states. *Philos. Sci.* 66, 175–207. doi: 10.1086/392683

Behrmann, M., and Plaut, D. C. (2013). Distributed circuits, not circumscribed centers, mediate visual recognition. *Trends Cogn. Sci.* 17, 210–219. doi: 10.1016/j.tics.2013.03.007

Bi, T., Cai, P., Zhou, T., and Fang, F. (2009). The effect of crowding on orientation-selective adaptation in human early visual cortex. *J. Vis.* 9, 13.1–10.

Borsboom, D., Cramer, A. O., and Kalis, A. (2019). Brain disorders? Not really: why network structures block reductionism in psychopathology research. *Behav. Brain Sci.* 42, 1–54.

Bouma, H. (1970). Interaction effects in parafoveal letter recognition. *Nature* 226, 177–178. doi: 10.1038/226177a0

Campbell, F. W., and Robson, J. G. (1968). Application of Fourier analysis to the visibility of gratings. *J. Physiol.* 197, 551–566. doi: 10.1113/jphysiol.1968.sp008574

Chen, J. Y., He, Y., Zhu, Z., Zhou, T., Peng, Y., Zhang, X., et al. (2014). Attention-dependent early cortical suppression contributes to crowding. *J. Neurosci.* 34, 10465–10474. doi: 10.1523/jneurosci.1140-14.2014

Chung, S. T., Li, R. W., and Levi, D. M. (2007). Crowding between first-and second-order letter stimuli in normal foveal and peripheral vision. *J. Vis.* 7, 10.1–10.

Coates, D. R., Levi, D. M., Touch, P., and Sabesan, R. (2018). Foveal crowding resolved. *Sci. Rep.* 8, 1–12.

Crutch, S. J., Schott, J. M., Rabinovici, G. D., Murray, M., Snowden, J. S., van der Flier, W. M., et al. (2017). Consensus classification of posterior cortical atrophy. *Alzheimer's Dement.* 13, 870–884.

Crutch, S. J., and Warrington, E. K. (2007). Foveal crowding in posterior cortical atrophy: a specific early-visual-processing deficit affecting word reading. *Cogn. Neuropsychol.*24, 843–866. doi: 10.1080/02643290701754240

Crutch, S. J., and Warrington, E. K. (2009). The relationship between visual crowding and letter confusability: towards an understanding of dyslexia in posterior cortical atrophy. *Cogn. Neuropsychol.* 26, 471–498. doi: 10.1080/02643290903465819

Daniel, P. M., and Whitteridge, D. (1961). The representation of the visual field on the cerebral cortex in monkeys. *J. Physiol.* 159, 203–221. doi: 10.1113/jphysiol.1961.sp006803

Falcinelli, G. (1993). "Osteo odontokerato prosthesis: 20 years of positive experience and innovations," in *Paper Read to Manhattan Eye*, New York, NY: Ear and Throat Hospital), 29.

Feigl, H. (1958). The 'mental'and the 'physical'. *Minnesota Stud. Philos. Sci.* 2, 370–497.

Flom, M. C., Heath, G. G., and Takahashi, E. (1963a). Contour interaction and visual resolution: contralateral efects. *Science* 142, 979–980. doi: 10.1126/science.142.3594.979

Flom, M. C., Weymouth, F. W., and Kahneman, D. (1963b). Visual resolution and contour interaction. *JOSA* 53, 1026–1032.

Fodor, J. A. (1974). Special sciences (or: the disunity of science as a working hypothesis). *Synthese* 28, 97–115. doi: 10.1007/bf00485230

Freeman, J., and Simoncelli, E. P. (2011). Metamers of the ventral stream. *Nat. Neurosci.* 14, 1195–1201. doi: 10.1038/nn.2889

Freeman, J., Donner, T. H., and Heeger, D. J. (2011). Inter-area correlations in the ventral visual pathway refect feature integration. *J. Vis.* 11:15. doi: 10.1167/11.4.15

Graham, N., and Nachmias, J. (1971). Detection of grating patterns containing two spatial frequencies: a comparison of single-channel and multiple-channels models. *Vis. Res.* 11, 251–259.

Grainger, J., Tydgat, I., and Isselé, J. (2010). Crowding affects letters and symbols differently. *J. Exp. Psychol.* 36, 673–688. doi: 10.1037/a0016888

Harrison, W. J., and Bex, P. J. (2017). Visual crowding is a combination of an increase of positional uncertainty, source confusion, and featural averaging. *Sci. Rep.* 7:45551. doi: 10.1038/srep45551

He, D., Wang, Y., and Fang, F. (2019). The critical role of V2 population receptive fields in visual orientation crowding. *Curr. Biol.* 29, 2229–2236.

Herzog, M. H., Sayim, B., Chicherov, V., and Manassi, M. (2015). Crowding, grouping, and object recognition: a matter of appearance. *J. Vis.* 15:5. doi: 10.1167/15.6.5

Hubel, D. H., and Wiesel, T. N. (1965). Receptive fields and functional architecture in two nonstriate visual areas (18 and 19) of the cat. *J. Neurophysiol.* 28, 229–289. doi: 10.1152/jn.1965.28.2.229

Hubel, D. H., and Wiesel, T. N. (1974). Sequence regularity and geometry of orientation columns in the monkey striate cortex. *J. Comp. Neurol.* 1, 267–293. doi: 10.1002/cne.901580304

Irvine, S. R. (1945). A review of solar retinitis as it may pertain to macular lesions seen in personnel of the Armed Forces. *Am. J. Ophthalmol.* 28, 1158–1165. doi: 10.1016/0002-9394(45)90215-7

Kas, A., De Souza, L. C., Samri, D., Bartolomeo, P., Lacomblez, L., Kalafat, M., et al. (2011). Neural correlates of cognitive impairment in posterior cortical atrophy. *Brain* 134, 1464–1478. doi: 10.1093/brain/awr055

Kay, K. N., and Yeatman, J. D. (2017). Bottom-up and top-down computations in word-and face-selective cortex. *Elife*, 6:e22341.

Kiorpes, L., and Daw, N. (2018). Cortical correlates of amblyopia. *Vis. Neurosci.* 35:E016.

Kiorpes, L., Kiper, D. C., O'Keefe, L. P., Cavanaugh, J. R., and Movshon, J. A. (1998). Neuronal correlates of amblyopia in the visual cortex of macaque monkeys with experimental strabismus and anisometropia. *J. Neurosci.* 18, 6411–6424. doi: 10.1523/jneurosci.18-16-06411.1998

Korte, W. (1923). Über die Gestaltauffassung im indirekten Sehen. *Z. Psychol.* 93, 17–82.

Kwon, M., Bao, P., Millin, R., and Tjan, B. S. (2014). Radial-tangential anisotropy of crowding in the early visual areas. *J. Neurophysiol.* 112, 2413–2422. doi: 10.1152/jn.00476.2014

Kwon, M., and Liu, R. (2019). Linkage between retinal ganglion cell density and the nonuniform spatial integration across the visual field. *Proc. Natl. Acad. Sci. U.S.A.* 116, 3827–3836. doi: 10.1073/pnas.1817076116

Levi, D. M. (2008). Crowding - An essential bottleneck for object recognition: a mini-review. *Vis. Res.* 48, 635–654. doi: 10.1016/j.visres.2007.12.009

Liang, P., Wu, S., and Gu, F. (eds). (2016). "Neurobiological basis underlying neural information processing," in *An Introduction to Neural Information Processing* (Dordrecht: Springer), 14.

Manassi, M., and Whitney, D. (2018). Multi-level crowding and the paradox of object recognition in clutter. *Curr. Biol.* 28, R127–R133.

Martelli, M. (2001). *Visual Recovery After Long Term Deprivation in Humans (Unpublished Doctoral Dissertation)*. Ph.D. thesis, Sapienza University of Rome, Rome.

Martelli, M., Majaj, N. J., and Pelli, D. G. (2005). Are faces processed like words? A diagnostic test for recognition by parts. *J. Vis.* 5, 58–70.

Martelli, M. L., De Luca, M., Falcinelli, G., Gatti, G., Spinelli, D., and Zoccolotti, P. (2000). "Pattern of visual recovery after long term deprivation in humans," in *Investigative Ophtalmology & Visual Science*, Vol. 41 (Bethesda, MD: Assoc Research Vision Ophtalmology Inc), S49–S49.

Millin, R., Arman, A. C., Chung, S. T., and Tjan, B. S. (2013). Visual crowding in V1. *Cerebral Cortex* 24, 3107–3115. doi: 10.1093/cercor/bht159

Motter, B. C. (2006). Modulation of transient and sustained response components of V4 neurons by temporal crowding in flashed stimulus sequences. *J. Neurosci.* 26, 9683–9694. doi: 10.1523/jneurosci.5495-05.2006

Nandy, A. S., and Tjan, B. S. (2007). The nature of letter crowding as revealed by first-and second-order classification images. *J. Vis.* 7, 5.1–26.

Parkes, L., Lund, J., Angelucci, A., Solomon, J. A., and Morgan, M. (2001). Compulsory averaging of crowded orientation signals in human vision. *Nat. Neurosci.* 4, 739–744. doi: 10.1038/89532

Pelli, D., and Rosen, S. (2015). Supposing that crowding is compulsory grouping suggests a remarkably simple model for object recognition. *J. Vis.* 15, 1094–1094.

Pelli, D. G., Palomares, M., and Majaj, N. J. (2004). Crowding is unlike ordinary masking: distinguishing feature integration from detection. *J. Vis.* 4, 1136–1139.

Pelli, D. G., and Tillman, K. A. (2008). The uncrowded window of object recognition. *Nat. Neurosci.* 11, 1129–1135. doi: 10.1038/nn.2187

Pelli, D. G., Waugh, S. J., Martelli, M., Crutch, S. J., Primativo, S., Yong, K. X., et al. (2016). A clinical test for visual crowding. *F1000Res.* 5:81. doi: 10.12688/f1000research.7835.1

Pernu, T. (2017). "Elimination, not reduction: Lessons from the Research Domain Criteria (RDoC) and multiple realisation. Open peer commentary," in *Brain Disorders? Not Really: Why Network Structures Block Reductionism in Psychopathology Research*, eds D. Borsboom, A. O. Cramer, and A. Kalis (Cambridge, MA: Behavioral and Brain Sciences), 42.

Petersen, A., Vangkilde, S., Fabricius, C., Iversen, H. K., Delfi, T. S., and Starrfelt, R. (2016). Visual attention in posterior stroke and relations to alexia. *Neuropsychologia* 92, 79–89. doi: 10.1016/j.neuropsychologia.2016.02.029

Petersen, S. E., and Sporns, O. (2015). Brain networks and cognitive architectures. *Neuron* 88, 207–219. doi: 10.1016/j.neuron.2015.09.027

Place, U. T. (1956). Is the consciousness a pprocess of brain. *Br. J. Psychol.* 47, 44–50.

Polger, T. W., and Shapiro, L. A. (2016). *The Multiple Realization Book*. Oxford: Oxford University Press.

Putnam, H. (1967). "Psychological predicates," in *Art, Mind, and Religion*, eds W. H. Capitan and D. D. Merrill (Pittsburgh, PA: University of Pittsburgh Press), 37–48.

Raichle, M. E., MacLeod, A. M., Snyder, A. Z., Powers, W. J., Gusnard, D. A., and Shulman, G. L. (2001). A default mode of brain function. *Proc. Natl. Acad. Sci. U.S.A.* 98, 676–682.

Rosen, S., Chakravarthi, R., and Pelli, D. G. (2014). The Bouma law of crowding, revised: critical spacing is equal across parts, not objects. *J. Vis.* 14:10. doi: 10.1167/14.6.10

Sand, K., Robotham, R. J., Martelli, M., and Starrfelt, R. (2018). Visual crowding in pure alexia and acquired prosopagnosia. *Cogn. Neuropsychol.* 35, 361–370. doi: 10.1080/02643294.2018.1483325

Shapiro, L. A. (2004). "The mind incarnate," in *Life and Mind: Philosophical Issues in Biology and Psychology*, eds K. Sterelny and R. Wilson (Cambridge, MA: MIT Press).

Shooner, C., Hallum, L. E., Kumbhani, R. D., Ziemba, C. M., Garcia-Marin, V., Kelly, J. G., et al. (2015). Population representation of visual information in areas V1 and V2 of amblyopic macaques. *Vis. Res.* 114, 56–67. doi: 10.1016/j.visres.2015.01.012

Smart, J. J. C. (1959). Sensations and brain processes. *Philos. Rev.* 68, 141–156.

Song, S., Levi, D. M., and Pelli, D. G. (2014). A double dissociation of the acuity and crowding limits to letter identification, and the promise of improved visual screening. *J. Vis.* 14:3. doi: 10.1167/14.5.3

Strappini, F., Galati, G., Martelli, M., Di Pace, E., and Pitzalis, S. (2017a). Perceptual integration and attention in human extrastriate cortex. *Sci. Rep.* 7, 1–15.

Strappini, F., Pelli, D. G., Di Pace, E., and Martelli, M. (2017b). Agnosic vision is like peripheral vision, which is limited by crowding. *Cortex* 89, 135–155. doi: 10.1016/j.cortex.2017.01.012

Treisman, A., and Schmidt, H. (1982). Illusory conjunctions in the perception of objects. *Cogn. Psychol.* 14, 107–141. doi: 10.1016/0010-0285(82)90006-8

Treisman, A. M., and Gelade, G. (1980). A feature-integration theory of attention. *Cogn. Psychol.* 12, 97–136. doi: 10.1016/0010-0285(80)90005-5

Tyler, C. W., and Likova, L. T. (2007). Crowding: a neuroanalytic approach. *J. Vis.* 7, 16.1–9.

Wandell, B. A. (1995). *Foundations of Vision*. Sunderland, MA: Sinauer.

Wandell, B. A., and Winawer, J. (2011). Imaging retinotopic maps in the human brain. *Vis. Res.* 51, 718–737. doi: 10.1016/j.visres.2010.08.004

Whitney, D., and Levi, D. M. (2011). Visual crowding: a fundamental limit on conscious perception and object recognition. *Trends Cogn. Sci.* 15, 160–168. doi: 10.1016/j.tics.2011.02.005

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: 04 September 2020
doi: 10.3389/fpsyg.2020.02058

Phenomenological Skepticism Reconsidered: A Husserlian Answer to Dennett's Challenge

*Jaakko Belt**

Unit of History, Philosophy and Literature, Faculty of Social Sciences (SOC), Tampere University, Tampere, Finland

There is a long-standing tradition of questioning the viability and scientificity of first-person methods. Husserlian reflective methodology, in particular, has been challenged on the basis of its perceived inability to meet the standards of objectivity and reliability, leading to what has been called "phenomenological skepticism" (Roy, 2007). In this article, I reassess this line of objection by outlining Daniel C. Dennett's empirically driven skepticism and reconstructing his methodological arguments against Husserlian phenomenology. His ensuing phenomenological skepticism is divided into strong skepticism and categorical and gradual versions of weak skepticism. Both strands of Dennett's criticism are then countered by analyzing the key components of Husserl's method of phenomenological reflection: *epoché* and transcendental reduction, intentional analysis, eidetic variation, and intersubjective validation. Laying out the basic features of phenomenological reflection serves two purposes. First, it undermines Dennett's methodological arguments, which are based on the unfounded assumptions that Husserl is committed to introspection, methodological solipsism, the first-person-plural presumption, and the lone-wolf approach. Second, it shows how Husserl's own methodology can alleviate the more justified empirical worries concerning overinterpretation, underdescription, and disagreement. Finally, I argue that gradual weak skepticism is the only plausible form of phenomenological skepticism and conclude that Husserlian methodology is well-equipped to combat it.

Keywords: Husserl, Dennett, phenomenology, methodology, reflection, introspection, skepticism, first-person methods

OPEN ACCESS

Edited by:
Christopher Gutland, Sun Yat-sen University, China

Reviewed by:
Jean-Michel Roy, Université de Lyon, France
Christian Tewes, Heidelberg University Hospital, Germany

***Correspondence:**
Jaakko Belt
jaakko.belt@tuni.fi;
jaakko.belt@gmail.com

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

Received: *08 February 2020*
Accepted: *24 July 2020*
Published: *04 September 2020*

Citation:
Belt J (2020) Phenomenological Skepticism Reconsidered: A Husserlian Answer to Dennett's Challenge. Front. Psychol. 11:2058. doi: 10.3389/fpsyg.2020.02058

INTRODUCTION

There is a long tradition of questioning the viability and scientificity of reflective methodology. Back in the early nineteenth century, Auguste Comte stated in his *Cours de philosophie positive* (1830–1842):

> "For all the two thousand years during which metaphysicians have thus cultivated psychology, they are not agreed about one intelligible and established proposition. [. . .] '*Internal observation*' gives almost as many divergent results as there are individuals who think they practice it."[1]

To capture this still prevalent concern, Jean-Michel Roy has coined the term "phenomenological skepticism." He defines phenomenological skepticism in terms of the current debate about the

[1]Cited in James (1890/1950, p. 188).

relevance of phenomenology, and first-person methods in general, in the context of cognitive science and the science of consciousness:

> "[P]henomenological scepticism [is] the long standing objection that the traditional conception of phenomenology falls short of the basic requirements of science, because it cannot provide a knowledge endowed with a sufficient degree of reliability and objectivity."[2]

It is no surprise that Roy focuses on Daniel C. Dennett as the most ardent present-day proponent of phenomenological skepticism. Much like other philosophical behaviorists before him, Dennett (2001/2018, p. 467) has voiced his distrust of first-person investigations, neither mincing his words nor lacking in rhetorical flair: "First-person science of consciousness is a discipline with no methods, no data, no results, no future, no promise. It will remain a fantasy."

Such an unequivocal condemnation of first-person approaches has led some commentators to argue that Dennett has construed an overly simplistic picture of first-person methods in general and a straw man view of classical phenomenology in particular (see Zahavi, 2007; Cerbone, 2012). Conflating different ways of examining conscious experience from the first-person perspective, especially introspection and phenomenological reflection, has allegedly led Dennett to think that their flaws and limitations are similar, if not the same altogether. His undifferentiated view of first-person methods and shared mistrust of them can be seen as Dennett's reason for rejecting all kinds of reflective endeavors, including Husserl's phenomenology, as scientifically suspect.

Roy (2007) claims, however, that Dennett's phenomenological skepticism should not be considered a wholesale dismissal of reflective efforts in general or even phenomenology in particular. On closer inspection, he maintains, Dennett employs a dual strategy of making use of some elements of Husserlian phenomenological methodology, while doing away with the rest[3]. On this reading, Dennett criticizes phenomenological reflection only on epistemological and methodological grounds, questioning its reliability and alleged lack of objectivity. At the same time, he is open to integrating Husserlian analyses into a naturalistic framework – in particular, into his "heterophenomenological" approach to the study of consciousness. This reading is supported by Dennett's self-avowed "buffet approach" to Husserl[4].

Those who regard Dennett as largely dismissive of phenomenology have also dissected his brand of skepticism. David Cerbone has reconstructed epistemological and ontological varieties of skepticism and offered detailed responses to ensuing questions concerning the accuracy and comparability of reflective knowledge. Dan Zahavi, in turn, takes on Dennett's charge of methodological solipsism by highlighting that structures of experience are intersubjectively accessible objects of reflective investigation and that phenomenological analyses result in descriptions and arguments open to communal corrections. In answering Dennett's critique, both Cerbone and Zahavi present Husserl's phenomenology as a transcendental project that investigates the conditions of experience and the constitution of reality. From the transcendental perspective, it is possible to turn the tables on Dennett and question his commitments to the naturalization of consciousness and objectivistic scientific worldview[5]. Both Zahavi and Cerbone also argue that Dennett mistakes phenomenological reflection for (psychological) introspection and deem skepticism deriving from this equation misguided[6]. In contrast, Roy (2007) recognizes phenomenological skepticism as a pertinent problem – especially if one wishes to integrate Husserlian investigations into a naturalistic framework of cognitive science. In his mind, Husserl's anti-naturalist *credo* and alleged commitment to infallibilism render "orthodox" Husserlian phenomenology vulnerable to Dennett's objection.

I will draw from Zahavi's and Cerbone's arguments and develop some of them further in order to counter Dennett's methodological criticism of Husserlian phenomenology. In taking on Dennett's empirically oriented arguments as objections worthy of closer consideration, however, my strategy is closer to Roy and his integrative approach, although Zahavi's and Cerbone's reading of Husserl is more faithful to his work. It is true that Husserl saw all genuine skepticism (including doubting the epistemic value of reflection) as self-defeating and countersensical (*widersinnig*), since it implicitly assumes or makes use of what it explicitly denies (Husserl, 1976, p. 174; cf. Husserl, 1975, p. 120, 123). At the same time, he was well aware of the same kind of worries associated with casual reflection and psychological introspection that also motivate Dennett's skepticism. Husserl (1976, p. 172) recognized that skeptical doubts concerning self-observation can be readily extended to all reflection, including phenomenological reflection. For this reason, I argue, Husserl's phenomenological reflection is well-equipped to safeguard against them. In this respect, my conclusion also differs from Roy's: rather than succumbing to phenomenological skepticism, Husserl's reflective methodology offers tools for

[2] Roy (2007, p. 4, cf. 9). The term itself is somewhat misleading, since skepticism in question is neither *based on* phenomenological findings nor *driven by* phenomenological attitude (in the sense that it is customary to speak of Pyrrhonian skepticism or Cartesian doubt, for instance). In contrast, it is skepticism *about* first-person methodology and its ability to reliably describe what Roy calls "phenomenological properties" and what Dennett (1991, p. 45) dubs as "phenomenology" (with a lowercase p), namely "the various items in conscious experience that have to be explained."

[3] Carr (1998) offers a more pessimistic view: due to lack of knowledge of phenomenological tradition and insufficient understanding of Husserlian methodology, Dennett ends up adopting elements of Husserl's reflective methodology without acknowledging it.

[4] See Dennett (2007, p. 267): "I am happy not just to concede but to insist that many of the brilliant reflections of Husserl and Husserlians ought to be exploited to the full in heterophenomenological research. I just want to strip them of the anti-naturalistic ideology that has – for the most part – weighed them down [. . .] we can salvage all the good ideas of Phenomenology and incorporate them into heterophenomenology." Cf. Dennett, 1978, p. 184.

[5] For more detailed arguments, see Zahavi (2007, 2017, p. 10–11, 16–17, 25–27, 123–126, 131, 144–150, 163); Cerbone (2003, 2012).

[6] Zahavi (2007, 2017, p. 9, 14–15, 27) and Cerbone (2003, 2012) provide both textual evidence and methodological reasons as to why Husserl adamantly distinguished his method from introspection. By contrast, Gutland (2018) argues that Husserl's methodology can be seen as a refined and systematized kind of introspection, which avoids some of its shortcomings.

restraining our erroneous tendencies and mitigating the skeptical concerns.

I will begin by outlining Dennett's empirically driven reasons for phenomenological skepticism (section "Dennett's Empirical Arguments"). Then, I will reconstruct his methodological arguments against Husserlian reflective methodology (section "Dennett's Methodological Arguments"). The focus will be on four methodological commitments Husserl allegedly makes: use of introspection, methodological solipsism, the first-person-plural presumption, and the lone-wolf approach. In section "Strong and Weak Skepticism," I will take a closer look at the ensuing phenomenological skepticism by dividing it into strong skepticism, categorical weak skepticism, and gradual weak skepticism. In order to respond to phenomenological skepticism, I will explicate the basic elements of Husserl's method of phenomenological reflection, namely *epoché* and transcendental reduction, intentional analysis, eidetic variation, and intersubjective validation (section "Phenomenological Reflection"). This has a dual function of dispelling Dennett's methodological arguments while showing how Husserl's methodology can alleviate the more justified empirical worries concerning overinterpretation, underdescription, and disagreement. Finally, I will conclude that gradual weak skepticism is the only plausible form of phenomenological skepticism and that Husserlian methodology is well-positioned to combat it.

DENNETT'S EMPIRICAL ARGUMENTS

Dennett's critique of first-person methods can be divided into empirical arguments and principled methodological worries. First, it should be noted that Dennett neither categorically denies the possibility of reflective description of conscious experience nor questions that subjects have *some* privileged access to their own experience. The problem is, rather, that no refined technique providing reliable results and methodological guidelines for scientific use of first-person methods has been established. In other words, Dennett argues that subjective approaches to studying conscious experience have failed to meet the epistemological standards and methodological requirements of science. Phenomenology, in particular, has not met its goal of describing the contents of our conscious experience faithfully and reliably, without distortions or unfounded theorizing. By lacking a neutral method of description and a common ground for assessing its results, Dennett claims that phenomenology "has failed to find a single, settled method that everyone could agree upon" (Dennett, 1991, p. 44, 67–69). I will challenge Dennett's assessment of methodological unanimity in section "Phenomenological Reflection." Let us first break down the empirical and methodological reasons for his suspicion of first-person methods.

Dennett points out several restrictions on our capacity to reflect upon conscious experience. In arguing against the viability of first-person investigations, though, he uses empirical research unsystematically and sporadically at best. To support his case, Dennett also resorts to traditional philosophical arguments and utilizes illustrations from the history of philosophy and everyday psychology, supplemented with analogies, metaphors, and thought experiments. Nevertheless, I call the following arguments *empirical* because they either have some basis in actual empirical research or at least point to human psychological or cognitive tendencies and the limits of our reflective capacities that can in principle be empirically detected and tested in experimental settings[7].

Dennett's empirical arguments center around the claim that we are overconfident in our ability to 'get it right' when it comes to our own experience. This propensity comes to the fore in, at least, three forms: (1) underdescription, (2) disagreement, and (3) overinterpretation.

First, we seem to underestimate the blind spots of our reflective grasp of our conscious experience. Important features of ongoing experience go unnoticed and, in some cases, we seem to be demonstrably mistaken about them. Dennett's favorite example is peripheral vision. According to Dennett, naïve reflection makes us think that our visual field is sharp and uniformly detailed not only from the center but also all the way to the boundaries. But even simple demonstrations (such as moving a playing card held at arms length from your side to the center of your visual field) show that, in fact, it is hard to identify objects in terms of their color or shape quite close to the center, even though you can detect movement. This so-called deficiency in our peripheral vision goes unnoticed because our eyes are normally continuously tracking and saccading in order to bring objects to the center of foveal vision. Instead of providing information in a manner of a "snapshot," our visual field is much more undetermined and lacking in detailed visual content, with only a rapidly shifting clear center. This finding is said to surprise most people, even cognitive scientists, to the effect that in test settings many subjects confess being formerly mistaken about their visual field[8].

Second, Dennett points out, in the Comtean vein, that factual disagreement and the lack of comprehensive data provide evidence for the *un*reliability of first-person methods. Dennett refers in passing to some examples from the history of empirical introspective psychology, like the unresolved debate about imageless thought (Dennett, 1991, p. 59; cf. Roy, 2007). His overall argument, however, does not rest solely on documented cases of disagreement. He also invokes the alleged lack of *positive* results of first-person investigations and the supposed inability to settle disputes if conflicts ensue. According to Dennett (1991, p. 44–45, 96; cf. Dennett, 1978, p. 185), phenomenologists, in particular, have failed to produce a catalog

[7]In the introspection debate, Peels (2016) defines empirical (scientific) arguments as arguments "somehow based on empirical scientific research." He evaluates five main arguments against the reliability of introspection in the current literature by looking case-by-case at whether the conducted experiments actually support the conclusion. In this article, my goal is simply to reconstruct the empirically motivated general arguments and lines of thinking behind Dennett's phenomenological skepticism, not to evaluate the actual or potential empirical support for his claims.

[8]Dennett (1991, p. 53–54, 68, 2001/2018); cf. Schwitzgebel (2011, p. 125–126). In empirical research, the phenomenon is often divided into distinct but closely related forms of inattentional blindness and change blindness (see again Peels, 2016).

of all the items inhabiting our conscious experience, whose contents the experts could by and large agree upon. Instead of being a reliable communal activity of "pooling shared observations," first-person investigations have allegedly lapsed into "the battle of 'intuitions'," where controversies are often met with "desk-thumbing cacophony" and "talking past everybody else" (Dennett, 1991, p. 66, 96).

Third, according to Dennett (1991, p. 67), this kind of "controversy and contradiction" – *contra* the sought-after mutual agreement – not only shows that our trust in high reliability of introspection is misguided; it also betrays the fact that we are prone to overinterpretation and unfounded theorizing about our experience. Dennett's notorious example is the notion of the self, which he sees as a narrative creation. It may be a useful fiction, but the self nevertheless is something neither reliably found within our conscious experience nor verified by external observation (Dennett, 1991, p. 412; cf. Dennett, 1992). It is highly unlikely that a question as multifaceted as the nature and existence of the self could ever be settled in an empirical setting. Nevertheless, Dennett thinks the impasse of modern philosophy is telling: the chain of philosophers who all claim to be using a first-person method of some kind and assume that their "introspecting" could be readily replicated at will (Descartes, Hume, Locke, and their successors), have ended up in conflicting, and even opposing, views on whether there is a self at all and what its nature would be (see Dennett, 1991, p. 66–67, 412–413). The variety of opinions indicates that humans have a tendency to fabulate descriptions of their own experience – to "fill in the gaps, guess, speculate, mistake theorizing for observing" (Dennett, 1991, p. 94)[9].

More recently, Eric Schwitzgebel has suggested three types of argument, all based on empirical case studies and their philosophical analysis, as to why introspection is prone to error: (1) There seems to be more variation in people's introspective reports than is plausible to assume there are underlying differences in ways we experience things (argument from variation)[10]. (2) There are cases where most people are clearly and crudely mistaken about quite basic features of their own experience (argument from error)[11]. (3) Sometimes introspection yields remarkably inconclusive results (argument from uncertainty)[12]. One can notice the overlap with Dennett's arguments; in fact, Dennett can be seen to apply these argumentative strategies in undermining our confidence in the reliability of reflective investigations.

I will return to the empirical arguments when assessing Dennett's phenomenological skepticism in section "Strong and Weak Skepticism." It is, however, already important to note that the restrictions detailed above are not specific to phenomenological reflection *per se*. The implicit argument found in Dennett is, rather, that phenomenology suffers from the same kinds of empirical limitations and misgivings as everyday reflection and earlier introspective psychology, until proven otherwise. In Dennett's view, we are bound to overstep, or to ignore altogether, the limits of our reflective cognitive capacities. In section "Phenomenological Reflection," I will formulate a Husserlian response to the challenges posed by Dennett's empirical arguments by showing how phenomenological reflection actually safeguards against the perceived problematic tendencies of casual reflection and introspection and offers methodological tools for alleviating overinterpretation, underdescription, and disagreement.

DENNETT'S METHODOLOGICAL ARGUMENTS

Empirical arguments in the wide sense outlined above can be separated from *methodological* arguments, which are based on Dennett's general view on what counts as science and what scientific methodologies allegedly permit. Dennett's methodological critique of phenomenological reflection is based on the distinction between first-person and third-person methods and their respective data[13]. In *Consciousness Explained*, Dennett makes a categorical claim that scientific theories can only be constructed from the third-person perspective. For Dennett, this means using objective methodologies that rely only on data that is intersubjectively accessible and verifiable, i.e., available for external observation and open for independent validation. In order to be considered as scientific, first-person investigations should also be able to constitute a reliable communal practice based on shared observations. But instead of delivering identical results or even findings that *could* be replicated, first-person investigations arguably end up relying on private access to subjective conscious experience and produce indefeasible statements (see Dennett, 1991, p. 66, 70–71, 2003, 2007; see also Overgaard et al., 2008).

Dennett sees his methodological criticism as resting on a standard conception of science and does not admit any need to reform or even adjust it to accommodate first-person methods. Therefore, he builds his case by pointing out *deviations* from the presumably widely accepted standards of natural sciences, rather than specifying criteria for scientific objectivity or reliability more explicitly[14]. By way of approximating these criteria, first-person methods fail to meet at least three underlying requirements of

[9]The well-known case study by Nisbett and Wilson (1977) has been taken to show that we, in fact, retrospectively posit, rather than introspectively observe, some key factors in our experience, even when we think we are only consciously reflecting. In Dennett's view, however, our tendency to fabulate extends well beyond the causes of our experiences and the self-attribution of reasons for action examined by Nisbett and Wilson (see Roy, 2007).

[10]As Schwitzgebel (2011, p. xi) notes, in addition to cross-cultural variation in reports, differences are found between individuals in the same cultural context *and* within one and the same individual over time. Dennett seems to concur: it is safer to assume that we are all more or less alike but that we err in describing our experiences because of the unreliability of introspection (cf. Dennett, 1991, p. 67).

[11]Schwitzgebel (2011, p. 126) also brings up peripheral vision as a case where untutored reflectors are, at first, usually badly mistaken, but can acknowledge their error when warned and trained.

[12]Schwitzgebel (2011, Ch. 1) points out that the question "Do we dream in color?" has produced vastly different answers in different times, suggesting that we may not be as firmly convinced as we think even on a subject that should be almost trivial. In Dennett's case, referring to the unresolved debate about imageless thought in introspective psychology might qualify as a sketch for an argument from uncertainty.

[13]Here, I follow Piccinini's (2010) reading of Dennett.

[14]Conversely, Dennett (2003, 2007) argues for his version of *hetero*phenomenology as an extension of science by stressing its conformity to the objective standards of

science: (1) they depend on a single observer and use private data as evidence (rather than producing intersubjectively testable statements on publicly available objects of investigation); (2) they rely on unreliable introspective practices (rather than reliable observation producing mostly accurate and corrigible results); (3) they provoke disputes about the methods and disagreement about their results (rather than producing unanimity and ways of settling the disputes). Let us call these criteria intersubjectivity or publicity, reliability, and agreement[15].

What reasons does Dennett give for his methodological concerns? In addition to citing reflective errors and empirical restrictions examined in the previous section, he turns to the history of psychology for indirect support of his position. Dennett sees the decline of introspective psychology and the resulting advancement of behaviorism in the first half of the twentieth century as bearing witness to the insurmountable problems of first-person methodology in general. In response to introspective psychology's inability to compare, replicate, and validate its results, behaviorists accepted only intersubjectively verifiable methods and dismissed any purported facts about mental events as data: since one cannot "look directly into" the minds of others, one should stick strictly to observation "from the outside." What is central to Dennett's argument is that the ensuing methodological shift was not restricted to the behaviorist school or confined to a certain period. Rather, in *Consciousness Explained*, he argues that suspicion of first-person methods has since become the guiding norm of *all* research in experimental psychology and neuroscience[16]. Dennett then proceeds to insist that a theory of consciousness *must* be constructed from the third-person perspective, "using the data that scientific method permits" and "never abandoning the methodological scruples of science" (Dennett, 1991, p. 71–72).

The methodological critique is also aimed at phenomenology more directly. By adopting what Dennett takes to be the standard first-person perspective in writing about consciousness, phenomenologists allegedly buy into what he calls the "first-person-plural presumption." By the *first-person-plural perspective* Dennett alludes to the resulting mode of investigation or a style of philosophizing where one describes first-hand in a monolog what is given in *my* conscious experience *and* assumes that others will agree. While Descartes' meditations and the first-person accounts favored by the British Empiricists are Dennett's paradigmatic historical examples of the approach, he sees modern day phenomenologists proceeding in a similar fashion. The way I read Dennett's criticism is that adopting the first-person-plural perspective leads to methodologically dubious practices in at least two related ways. First, it promotes problematic, and even careless, ways of generalizing from *my own* singular experience. Second, it paints a simplistic, and even erroneous, picture of first-person investigations as an agreement-producing practice, in which first-person accounts are readily reproduced by *others* by making personal inner observations and arriving at the same results (Dennett, 1991, p. 66–67, 70).

Dennett does not claim that first-person investigations are useless, not to even mention impossible. He admits that they may very well offer motivation, illustrations, and even guidance for scientific theories. But in his mind, they do not yet provide the kind of reliable data needed for the science of consciousness. First-person investigations become scientific only after the private reflective findings are turned into intersubjectively accessible third-person data by conducting controlled experiments with naïve test subjects. Dennett sees classical phenomenology committing itself to a "lone-wolf" approach, where both the subject and the object of investigation are one and the same person. By Dennett's standards, this puts phenomenologists in a similar position with experimenters who would run pilot studies on themselves but fail to confirm their findings with other test subjects. What Dennett seems to imply is that lone-wolf phenomenologists do not even try to meet the obligation of testing their insights intersubjectively and in interaction with others. Instead, they rely on "personal introspection" as the only evidence needed for substantiating claims about conscious experience. For Dennett, it is widely accepted that no defensible first-person *science* can be built on these grounds (Dennett, 2003).

In addition to the first-person-plural presumption and the lone-wolf approach, Dennett sees Husserlian phenomenology as married to a third methodological commitment, namely methodological solipsism. Dennett's methodological argument is ambiguous on this score. On the one hand, by methodological solipsism he means adopting *a research strategy* in which the experiencing subject is investigated in isolation from their environment, that is, not historically, linguistically, or causally embedded in the world[17]. On the other hand, it concerns different ways of *implementing* this strategy, that is, the methods used in gaining access to and studying the proper domain of investigation dictated by methodological solipsism. For the time being, it is sufficient to say that Dennett seems to think that an introspective approach combined with a methodological procedure called *epoché* or "bracketing" leads Husserl to exclude the outer world

natural sciences and its continuity with existing (reliable) practices in cognitive science, psychology, and scientific study of consciousness. He seemingly consents to his critics in saying that heterophenomenology could have easily been *labeled* as "the second-person method of gathering data" or even "first-person science of consciousness" (Dennett, 2007, p. 252, cf. 263–264). But what are the implications for objective third-person science if it is founded on interpersonal research practices, i.e., studies conducted from the embodied first-person or second-person perspective (or based on intersubjectively constituted life-world, to use Husserlian terms)? Such questions do not prompt Dennett to reform his conception of standard science or to refine the dichotomy of first-person and third-person perspective.

[15]Goldman (1997) traces such requirements back to a long line of philosophers of science defending the publicity of science as the core principle of any scientific methodology from Boyle to the positivists, Popper, and Hempel.

[16]Dennett (1991, p. 70). The skeptical consensus against first-person accounts was not as unified as Dennett's testimony of that time might suggest. See Velmans (1991) for the then ongoing debate on the mutual irreducibility and complementarity of the first-person and third-person perspectives in studying consciousness (and related phenomena) amongst empirically oriented psychologists and cognitive scientists. For a recent methodological discussion in the interdisciplinary field of integrative human neuroscience, see Kotchoubey et al. (2016).

[17]Dennett (1987, p. 134–135, 141, 153–154). Dennett adopts the term "methodological solipsism" from Hilary Putnam but in *The Intentional Stance* he focuses on Fodor's (1980) formulation of it. For the perceived affinities between Husserl and Fodor as methodological solipsists, see Dreyfus (1980), McIntyre (1988).

and its real objects from phenomenological reflection and to restrict its scope to the internal or mental domain and the subjective world of a person.

Supplementing this threefold negative critique, Dennett offers a more positive vision on integrating certain kinds of phenomenological investigation into a naturalistic framework. Instead of examining conscious experiences directly as they appear to a single subject living through them, Dennett suggests that the science of consciousness should study (other) people's *beliefs* about their experiences and their *reports* expressing them in experimental setting. He famously labels this methodological change of focus as a transition from *auto*phenomenology (phenomenology of oneself) to *hetero*phenomenology (phenomenology of the other). The researcher should not introspect, describe, and "catalog" conscious experiences as the sole reflector. Instead, she should gather first-person perspective reports of what people think of their experiences and interpret them by adopting the intentional stance[18]. Thus, what becomes "the data" is not the experiences themselves but people's beliefs *about* them, expressed in verbal form or through behavioral manifestations, like pushing a button or reacting in certain ways in different circumstances: "the reports *are* the *data*, they are not reports *of* the data" (Dennett, 1993; see also Dennett, 2003). Dennett (1991, p. 72) presents heterophenomenology as a neutral method of capturing even "the most private and ineffable" experiences without leaving the framework of objective third-person science and its standard methodologies. The picture of an armchair philosopher meditating in solitude is replaced with a figure of an empirical researcher gathering factual data and interpreting it with anthropological, psychological, and narrative insight.

STRONG AND WEAK SKEPTICISM

Dennett's phenomenological skepticism is based on the empirical and methodological arguments outlined above. At the outset, Dennett does not espouse strong skepticism, which claims that reflection has no access to anything real at all. He does not wish to prejudge the issue by simply denying the existence of conscious experience (or mental states) or identifying them with something else, e.g., information-bearing brain events[19]. Instead, Dennett's aim is to create a neutral method of describing experiences that is not committed to any ontological theses on the existence of experiences. This is why he presents his heterophenomenology as a method that investigates conscious experience indirectly *via* interpreting people's reports. In this way, Dennett hopes to take seriously how things *seem* or *appear* to subjects (i.e., what it is like to them), but grant them neither infallibility nor a final say on the fact of the matter (i.e., what is going on in them). This would arguably help to constrain unfounded theorizing about the nature and metaphysical status of objects of investigation and their assumed causes. For Dennett, commitment to neutrality is needed not only to secure a reliable way of gathering data, that is, for mapping out and describing experiences carefully, accurately, and comprehensively; a neutral way of extracting a "heterophenomenological catalog" also provides shared ground for settling reflective disputes, in order to avoid "the battle of 'intuitions'" allegedly characteristic of first-person investigations (Dennett, 1991, p. 96).

Dennett's plea for ontological neutrality comes with what he calls "metaphysical minimalism" (Dennett, 1991, p. 95). According to this principle, the objects of heterophenomenological investigation should not be taken at face value and be accepted as real, but taken only as assumptions, "theorists' fictions." Dennett leaves it to the empirical sciences to decide whether items described by heterophenomenology correspond to anything real or not, that is, whether they exist as real objects (as brain states, mental states, cognitive processes or the like) (Dennett, 1991, p. 81, 96, 98). Zahavi (2007) has claimed that here Dennett is actually an eliminativist in disguise, since his "metaphysical minimalism," in fact, leads to the denial of existence of experiences (cf. Cerbone, 2003, p. 133). Carr (1998) has pointed out problems with the fiction analog, starting with the fact that the metaphor runs counter to commitment to ontological neutrality, since *fiction* is not a metaphysically neutral term. In some cases, Dennett also seems to downright deny the actuality of experiences and in other cases he calls experiences "theoretical constructs" (see e.g., Dennett, 1991, p. 95, 157, 365). When Dennett (1991, p. 83) states that for a heterophenomenologist it makes no difference if her test subjects are "liars, zombies, or parrots dressed up in people suits," his ontological neutrality seems to lapse into what Cerbone (2012, p. 14) calls ontological skepticism: "it does not matter all that much whether anything really does correspond – whether there is a 'fact of the matter' about experience – since the beliefs are all that we have to express, report". More recently, Dennett has openly endorsed a strong anti-realist position called illusionism. Illusionism claims that phenomenal consciousness only *seems* to exist but is in fact illusory and that phenomenal properties are not real in the sense of being instantiated in the world (Frankish, 2016). For Dennett (2016), illusionism is not only the correct interpretation of his own account, but also the *default* view everyone engaging in scientific study of consciousness should hold, until proven otherwise. It is hard not to interpret these commitments as strong skepticism, well beyond the coveted ontologically neutral standpoint of heterophenomenology.

Dennett's general motivation for embracing such deflationist views can be traced back to his philosophical behaviorism, naturalistic approach to consciousness, and commitment to the third-person conception of standard science. Zawidzki has suggested that Dennett's strategy is to deflate *and* revise

[18]By the intentional stance, Dennett means a strategy of treating the object of investigation as a rational agent whose behavior can be predicted in terms of beliefs, desires, and other mental states characterized by intentionality or aboutness (see Dennett, 1987, p. 15, 24, 31–32, 1991, p. 76–77). While using this strategy is usually habitual and effortless, Dennett underlines that the intentional stance adopted freely in heterophenomenology differs from normal interpersonal relations: instead of trusting other subjects' word on their own experience, one is to maintain "constructive and sympathetic neutrality" (Dennett, 1991, p. 83) by neither challenging nor accepting the veracity of their assertions (see also Dennett, 1991, p. 81, 2001/2018; Zawidzki, 2018).

[19]Dennett (1991, p. 71, 459). By no means does Dennett shy away from developing and *eventually* adopting eliminativist and reductionist positions on certain issues. He simply does not want to embrace such views as a *starting point*. See Dennett (2007).

our manifest concepts of intentionality and consciousness in terms of publicly observable patterns of behavior. As intersubjectively verifiable phenomena, describable from the intentional stance and captured by heterophenomenology, they are, then, arguably easier to reconcile with the scientific image, i.e., what standard science accepts and rules out (Zawidzki, 2007, p. 154–156; cf. Dennett, 1987, p. 25). A case in point is Dennett's adamant rejection of *qualia* or phenomenal properties conceived as ineffable, intrinsic, private, systematically incomparable, and directly apprehensible features of conscious experience (Dennett, 1988; Frankish, 2016; cf. Dennett, 2020). Dennett denies the existence of *qualia* so defined as objectively undetectable, intersubjectively untestable, and, thus, incompatible with standard science. Hence, Dennett regards investigating consciousness scientifically in *these* terms impossible (see Zawidzki, 2007, p. 167–169). Whether Dennett's self-avowed eliminative materialism and verificationism about *qualia* amounts to denying conscious experience conceived in less contentious terms is debatable[20]. Whatever the case, Dennett has recently admitted that his battle against *qualia* as real properties of experience has made him overlook tensions in his discussion of conscious experience in terms of *seeming(s)*[21]. His ongoing polemics against "qualophiles" might, therefore, explain not only rhetorical excesses but also perceived ontological and epistemological inconsistencies in Dennett's views[22].

Perhaps a combination of deflationist-revisionist strategy and unyielding resistance to *qualia*, then, accounts for Dennett's apparent commitment to strong skepticism. In the next section, I will argue, however, that Dennett can defend his maxim of ontological neutrality and do away with *qualia* without espousing ontological skepticism, let alone illusionism or eliminativism about conscious experience. Therefore, I will focus on the weaker and more plausible forms of skepticism.

Weaker forms of skepticism only deny the possibility of (scientific) knowledge about our conscious experience. Dennett's epistemologically motivated skepticism rests on a conviction that we are far less immune to error than we think, a view that is supported by the types of empirical arguments outlined above. Dennett (2001/2018) refers specifically to two kinds of false beliefs, in which our assertions about our conscious experience fail to overlap with the fact of the matter: (i) false positives and (ii) false negatives. By *false positives* he means cases in which we believe we are conscious of more than what is in fact going on in us (as Dennett pointed out with the example of unsuspected deficiency in our peripheral vision). Conversely, by *false negatives* he refers to cases in which we do *not* believe we are conscious of things that are or were actually going on in us (as is demonstrated, for instance, by psychological experiments using masked priming, showing influence undetected from the first-person perspective). It is safe to assume that Dennett sees our inability to access all the workings of our own mind *via* "inner observation" as leading to both overstepping what can be reliably said about our experience (argument from overinterpretation) and overlooking important features of our experience (argument from underdescription) (see Dennett, 1991, p. 68–70, 94). He explicitly states false negatives and false positives as reasons for his skepticism about the truthfulness of first-person reports and as informing his insistence on bracketing the veracity of subjects' beliefs (methodological argument for ontological neutrality) (Dennett, 2001/2018, p. 457–458; cf. Dennett, 1991, p. 94).

It should be noted, however, that issues of false positives and false negatives reveal more about the restricted scope of reflection and uncertainty of its results than about the impossibility of first-person knowledge in general. It does not follow from the fact that some factors or features of our experience are not directly accessible from the first-person perspective that *all* of consciousness is inaccessible. Similarly, the inaccuracy of some reports does not merit the conclusion that conscious experience cannot be faithfully described at all. After all, one does not have to treat all reflective descriptions as equally reliable or assume that our experience is fully transparent to us (Goldman, 1997, p. 529, 532; Roy, 2007, p. 14). The problem is not that we can know *nothing* about our experience, but that we are overstating what we know (cf. Nisbett and Wilson, 1977). Another way to put it is to say that we take too large a portion of our mental and cognitive processes to be available reflectively and think that conscious experience is more transparent than it really is.

Following Hurlburt and Schwitzgebel (2007, p. 27), it might be helpful to distinguish skepticism about first-person access to non-conscious processes (such as cognitive mechanisms, but also psychological traits, behavioral dispositions, etc.) from skepticism about the self-reports of conscious experience. The defender of reflective knowledge can readily admit that the limits of privileged access are not always clear and that we might occasionally "overstep" even the "self-imposed restraints" (Dennett, 1991, p. 68) to stick within those limits. Fallibilism in these regards does not merit general skepticism about *all* possible reflective knowledge. I call this the *gradual version of weak skepticism*. As I will argue in the next section, Husserl's phenomenological reflection is well-positioned to take these worries into account.

Dennett's skeptical argument against phenomenology, though, also comes in a more categorical form. Dennett (1991, p. 67) claims that the Cartesian tradition sees us as either infallible (i.e., always guaranteed to be right) or at least incorrigible (i.e., no one can correct us) when it comes to knowing our conscious experience *via* self-observation. He also speaks of a

[20] See Dennett (1988, 1991, p. 126, 132, 390, 403, 454–455, 459–462). In his debate with David Papineau (see Crane, 2017), Dennett wavers between eliminativist and revisionist formulations in motivating his project of naturalizing consciousness: "It has been my work for fifty years to provide alternative, naturalistic alternatives to all these chimeras; it is not that consciousness doesn't exist but that it isn't what you probably think it is. (Qualia are a theorist's illusion but the subjectivity of experience is real.)"

[21] Dennett now accepts a distinction he earlier disavowed, namely the dichotomy between the way things (actually or objectively) seem to you as opposed to how they *seem* to seem to you. See Dennett (2018) commenting on Rosenthal (2018), cf. Dennett (1991, p. 132).

[22] Schwitzgebel (2007) has pointed out that Dennett, for instance, claims there is no *actual* phenomenology while *not* wishing to deny that conscious experience is something real. He also holds that we are never infallible about our own conscious experience while granting us limited incorrigibility – the last word – on our reports how it *seems* to us or "what it is like to be us."

"long-standing philosophical tradition" guilty of assuming that "*we all agree* on what we find when we 'look inside' at our own phenomenology" (Dennett, 1991, p. 66). Here, Dennett seems to imply that some kind of immunity from error is inherent in the aforementioned first-person-plural presumption. As an extension of the principle of charity, Dennett does propose granting limited incorrigibility to test subjects in heterophenomenological settings. People might be unreliable guides in what is going on in them, but they should be taken at their word on *what it is like* to be them or how it *seems* for them. But if one, then, goes on to extrapolate from one's own particular experience or engages in theorizing about the nature of conscious experience, this privilege is renounced and the person reflecting must be treated as a fallible theorizer (Dennett, 2002). Following this view, appealing to unchallengeable first-person authority or "papal infallibility" (as Dennett likes to put it) in an attempt to secure reflective knowledge about the general features of conscious experience is both methodologically unwarranted and empirically suspect (if the empirical arguments are any indication of our liability to error). Discrediting reflective knowledge *in toto* by invoking the unfounded reliance on infallibility or incorrigibility extended beyond its methodological confines can be named the *categorical version of weak skepticism*.

It is important to note that Dennett's categorical skepticism does not hinge on the claim that autophenomenology always or even typically gets it wrong. The charge is, rather, that autophenomenology claims first-person authority, while not acknowledging that reflection is always *susceptible* to error (see Dennett, 2002, 2007). It is not perfectly clear whether Dennett sees Husserl as being guilty of this traditional tenet of infallibility, but this interpretation merits a closer look. In the debate about skepticism, commentators like Roy see Husserl as defending the indubitability of phenomenological reflection in general: "Husserl was a firm believer that the immediate knowledge at the source of the phenomenological inquiry, which he technically labeled immanent intuition, is immune to error" (Roy, 2007, p. 6).

One could also claim that Husserl did not see the scope of reflection as limited. At least in principle, Husserl argues that we can, at any time, turn our attentive regard to our ongoing experience and to what is "straighforwardly" experienced in it. In other words, whenever there is conscious experience, we are free to reflect upon it, thus creating a second-degree act of reflection, which takes the previously lived-through experience and its meaning-content as its object (see Husserl, 1950, §§14–15, 1976, §§77–78). One possible reading of the claim is that the potential range of reflection is equivalent to the domain of experience as a whole. Perhaps this is what led Roy (2007, p. 14) to argue that Husserl was committed to a "double thesis," according to which "the whole of the mental is manifested to us, and that it is manifested to us as it really is," namely without distortions or limitations.

If Husserl admits neither restrictions to the scope of reflection nor uncertainty of its results, does he not succumb to the kind of infallibility and incorrigibility Dennett's categorical weak skepticism argues against? In order to answer this question in the negative, let us next turn to the methodological elements of phenomenological reflection.

PHENOMENOLOGICAL REFLECTION

In this section, I will reconstruct the key features of Husserl's method of phenomenological reflection. The reflective methodology can be articulated by dividing it into four elements: (1) *epoché* and transcendental reduction, (2) intentional analysis, (3) eidetic variation, and (4) intersubjective validation.

This reconstruction has both a negative and a positive function. First, it shows that Husserl's reflective methodology is not committed to the kind of presumptions Dennett claims, while addressing the more justified skeptical worries and potential pitfalls associated with first-person methods. Second, it outlines phenomenological reflection as a systematic method in its own right. Contemporary efforts to develop Husserlian approaches to conscious experience and to integrate them with cognitive sciences have yielded numerous applications, including direct and indirect use of phenomenology in experimental settings[23]. Whether one employs Husserlian methodology as a toolbox for compound methods (Schmicking, 2010) or wishes to develop it further as a self-standing method, understanding the core features of phenomenological reflection is invaluable. Hence, the methodological and epistemological issues discussed in this section have further relevance to the phenomenologically informed cognitive science and the interdisciplinary study of consciousness.

Epoché and Transcendental Reduction

Already in *Logische Untersuchungen* (1900–1901), Husserl called "presuppositionlessness" a basic principle of phenomenological research. This *Voraussetzungslosigkeit* demands refraining from metaphysical presumptions, psychological presuppositions, and theories and explanations of other sciences (both empirical-inductive and axiomatic-deductive), in order to consider experiences faithfully (Husserl, 1984, p. 24–28). Furthermore, phenomenological reflection has to fight both habits and tendencies rooted in our psychological development and linguistic problems in describing experience and communicating the results (Husserl, 1984, p. 14–15). In other words, from early on, Husserl points out several potential sources of distortion and

[23]Gallagher (2003) distinguishes three such approaches: neurophenomenology, front-loaded phenomenology, and indirect phenomenology. For *neurophenomenology*, cognitive neuroscience and phenomenological analyses are related through mutual constraints and advanced reciprocally (Varela, 1996). This involves training both scientists and test subjects in phenomenological practice to elicit refined first-person descriptions and descriptive categories, which are then used to analyze the data correlating with third-person behavioral and brain activity measurements (Gallagher, 2003). Early neurophenomenological case studies already used open questions in formulating descriptive categories (Lutz, 2002), but second-person interview techniques have since been refined in the neurophenomenological context and further developed into a novel approach called *micro-phenomenology* (Petitmengin, 2006, 2010; Petitmengin et al., 2019). *Front-loaded phenomenology*, in turn, utilizes the analytical tools and insights of phenomenological research in designing experimental setups and interpreting their results, without having to rely on training test subjects or requiring introspective reports. Finally, *indirect phenomenology* applies phenomenological insights to (re)interpret experimental results retrospectively. As Gallagher notes, indirect use of phenomenology eventually reraises the questions of how to use the phenomenological method *directly* and how to incorporate its result into the experimental context.

bias in reflection and description, professing a need to constrain them methodologically.

Throughout his career, Husserl also carefully heeded how reflection modifies its object, i.e., lived experience. In fact, reflecting, as Husserl sees it, is far from a process of simply recording and reporting what is already there – a matter of merely "looking and seeing" as Dennett (1991, p. 55, 66) claims we tend to think. At first, we are confronted by "dumb" or "mute" (*stumm*) experience (Husserl, 1950, p. 77). The cognitive value of phenomenological reflection is in its power to modify the previously unregarded and straightforward pre-reflective experience, in order to thematize and articulate it[24]. Dennett *could* argue that it is precisely this modification that alters and potentially distorts experience, making its faithful description impossible and opening up unconstrained theorizing[25]. Indeed, Husserl would admit that it is impossible to repeat or replicate the original experience exactly as it was lived through. At the same time, he would stress that this is not the goal of phenomenological reflection in the first place. After all, when we are immersed in our everyday living and take interest in the environing world, our experience is often seemingly undifferentiated. It is the reflective attitude that makes it possible to uncover, analyze, describe, and even clarify experiences in terms of their constituents and basic structure. To be sure, Husserl is well aware of the risks of reflective modification leading to "metaphysical construction" (Husserl, 1950, p. 177) and "reading into [experiences] more than is purely seen" (Husserl, 1950, p. 74) that Dennett warns us about. It is precisely for this reason that Husserl (1950, p. 74) sees preserving the "unprejudicedness" (*Vorurteilslosigkeit*) of descriptions as essential for acquiring (critical) reflective knowledge.

How can we combat the aforementioned difficulties methodologically and pursue the epistemological potential of reflection without prejudice? It was not until *Ideen I* (1913) that Husserl supplemented his basic principles of phenomenological investigation with a systematically spelled out doctrine of the phenomenological reduction(s). At its core, the doctrine involves an element of abstaining from judgments and thus "bracketing" or neutralizing our prior commitments; this is what Husserl famously calls *epoché*. Husserl speaks of "bracketings" (as well as reductions) in plural and alludes to different steps of the process, indicating that the scope of the procedure in question can be gradually expanded and restricted. Husserl's motive for this mode of presentation is both didactical and critical: we need to be constantly reminded not to let premises from other sciences (not only natural sciences, psychology, and *Geisteswissenschaften* like history and cultural and social sciences, but also formal-eidetic sciences such as pure logic and mathematics) carry over to phenomenology and instructed not to make use of their results as readymade stocks of knowledge. In his mind, presenting the needed "bracketings" step-by-step protects methodologically against common misconceptions (both contemporary and historical) and prepares for the avoidance of the constant threat of categorical mistakes and other ingrained habits of thought (see Husserl, 1976, §1, §56, §59, §61).

In the pregnant sense, however, *epoché* amounts to more than a series of exclusions. Rather than suspending our beliefs one by one or domain after domain, the ultimate aim is to "put out of action" *all* positing that characterizes our so-called natural attitude and its "general thesis." In other words, *epoché* has a universal goal of putting our everyday doxic attitude toward the world on hold and bracketing the related ontological commitments concerning its objects (Husserl, 1976, §§30–32).

Dennett does recognize Husserl's *epoché* as a possible way forward for securing the neutrality of descriptions. He even presumes heterophenomenology applies a third-person analog to *epoché* in reserving the judgment about the veracity of subjects' beliefs and seeking theory-free descriptions (Dennett, 2003). Even though the aim of neutrality might be the same, Husserl's and Dennett's varieties of "bracketing" are, in practice, quite different. In fact, as Cerbone (2003, p. 111) has noted, their views are almost mirror images in this respect: whereas Husserl brackets the whole of reality *including* ourselves as part of the world, Dennett puts into brackets only the reality of consciousness. Dennett, thus, comes close to doubting (ontological skepticism) or even denying (illusionism, eliminativism) the existence of conscious experience without questioning his own commitment to ontological naturalism. Husserl (1976, §§31–32, §109) himself takes pains to ensure that *epoché* should *not* be understood as a methodological *doubt* of existence, let alone as *negating* the actuality of conscious experience and the world. The aim is rather "neutralizing" or "parenthesizing" our ontological and theoretical commitments in order to examine them.

Following Husserl's original idea of *epoché* more closely would arguably help Dennett to attain his goal of ontological neutrality, without lapsing into strong skepticism. Positing superfluous entities, such as *qualia*, as a stand-in for phenomenal properties of experience could be avoided while resisting the opposite pull of simply reducing or identifying conscious experience with something else (see Dennett, 1969/1986, p. 112–113, 1991, p. 459–460). In this way, one could salvage ontological neutrality as one of the main motivations for both classical and heterophenomenology by separating it from the more problematic principle of metaphysical minimalism. Furthermore, a more faithful understanding of *epoché* could relieve some of the concerns behind gradual versions of weak skepticism. As a preparatory stage of Husserl's reflective methodology, the negative aim of *epoché* is to secure the aforementioned "unprejudicedness" of descriptions by identifying and avoiding different sources of bias. At the very least, *epoché* would safeguard against the sort of "impromptu theorizing" and over-interpretation Dennett is worried about. This is only half of the story, but Dennett could take this line in order to incorporate the negative function of *epoché* into heterophenomenology instead of interpreting *epoché* as a form of methodological solipsism.

The other option is to argue that Dennett cannot achieve ontological neutrality at all without also reflecting upon the presuppositions and the metaphysical baggage of his own

[24] Husserl (1976, §78). Several commentators have pointed out that this transformative modification is instrumental in attaining reflective knowledge (see Thomasson, 2003; Zahavi, 2015).

[25] Husserl (1984, p. 15) himself acknowledges the same problem.

naturalistic commitments (see Cerbone, 2012). In fact, it is arguably paramount for any theory of consciousness to reflect upon its basic assumptions concerning the place and function of consciousness in what is taken to be objective reality[26]. In addition to the negative move of excluding prejudices, *epoché* enables a positive methodological step or a change of attitude that Husserl calls *transcendental reduction*. According to Husserl (1950, §15), this philosophical procedure enables a new kind of *transcendental* reflection since it focuses on the constitution of reality and its preconditions in conscious experience. Husserl contrasts transcendental reflection with natural or psychological reflection which takes its objects to be part of the mental or "inner" domain as opposed to physical reality or the outer world. Transcendental reflection opens up philosophical inquiry into the (inter)subjective sources of meaning-formation or sense bestowing, constituting our understanding of and our belief in what is real. In sum, a fully effected *epoché* conjoins the positive and negative aspects of the method by *opening* a field of (transcendental-)phenomenological research and *securing* it from intrusive influences (Husserl, 1976, §32, §61).

The transcendental argument against Dennett, and all naturalistic positions for that matter, maintains that explaining the relationship between consciousness and reality without reflecting upon how conscious experience shapes our sense of reality in the first place is question-begging (Zahavi, 2007, 2017, p. 144ff; Cerbone, 2012). Along these lines, it can be claimed that Dennett fails to appreciate the fundamental methodological role of *epoché* and the ensuing transcendental reduction. Rendering *epoché* simply a matter of reserving judgment, suspending presuppositions, and striving for a theory-neutral description of experience (see Dennett, 2003) might partly capture the negative function of bracketing. However, as Zahavi has argued, refraining from preconceptions, speculation, and causal explanation in favor of unprejudiced description is neither the novel feature of *epoché* nor its ultimate goal. Presuppositionlessness was already one of Husserl's core principles before his transcendental turn and the introduction of *epoché* as an explicit method. Focusing only on what is excluded would miss the positive side of universal *epoché* as bringing about a thorough change of attitude toward the whole of reality (Zahavi, 2002, p. 111–112, 2019).

However, Dennett *does* think that *époche*, as Husserl applies it, actually involves a methodological change in how the subject's relationship to their environment is to be analyzed. This is, after all, implied by his methodological argument that (auto)phenomenology is committed to methodological solipsism and the lone-wolf approach. According to Dennett's interpretation of Husserlian methodology, *epoché* serves as the way of implementing methodological solipsism by bracketing the outer or real world. Thus, it circumscribes the field of study and limits investigation to inner observation of a single subject. Consequently, phenomenological reduction amounts to nothing more than a dubious "introspectionist bit of mental gymnastics" leading to one's *own* notional world (Dennett, 1987, p. 153, 157–158, 161). In other words, Dennett does not settle for his more modest claim that the goal of *epoché* is neutrality in the general sense of putting our presuppositions on hold when describing experience. Instead, he adopts a more problematic line of (mis)interpretation by holding that *epoché* and the ensuing reduction are needed for reorienting our focus from worldly objects to one's own inner experience[27].

It seems fair to conclude that Dennett's reading of *epoché* (and reduction) is at best selective and partial. At worst, it distorts his overall conception of phenomenological reflection, especially regarding its relation to the world and to the single reflecting subject. It is not uncommon to treat *epoché* and reduction as an inward turn leading into a solitary individual consciousness; after all, this fuels the interpretation of Husserl as a methodological solipsist (see e.g., Varela et al., 1991/2016, p. 16–17). In the next subsections, I will show, however, that seeing reflection as a solipsistic and introspectionist technique leads Dennett to overlook other important aspects of Husserlian reflective methodology. Turning to Husserl's ideas of intentional analysis and eidetic variation not only demonstrates why Dennett's interpretation of Husserl as a methodological solipsist is wrongheaded; Dennett's failure to acknowledge these ineliminable features of Husserl's methodology also makes clear how his misconstrual of *epoché* in the end contributes to his claim that Husserlian phenomenology is a form of lone-wolf autophenomenology operating in the first-person-plural presumption.

Intentional Analysis

In weighing Dennett's case against Husserl, perhaps the most revealing aspect is what is missing from his charge of methodological solipsism. To my knowledge, Dennett does not discuss Husserl's analysis of intentionality in any detail, save for a passing reference to *hyle* (Dennett, 1978, p. 184, 333) and comments on *noemata* as "the pure objects of conscious experience" (Dennett, 1991, p. 44) or "intentional objects constituted on a personal level" (Dennett, 2007, p. 259). As far as I can see, the nearest he gets to closing in on Husserl's account of intentionality is in his allusions to Brentano's doctrine of intentionality. In fact, Dennett's charge of methodological solipsism can be dissected into mentalistic, internalist, and fictionalist components, and there are echoes of Brentano's doctrine in each of these characterizations. Therefore, it is worth briefly tracing the Brentanian elements in Dennett's reading, before presenting Husserl's own version of intentional analysis as

[26]One could argue that in the last three decades Dennett has indeed reflected on the place and function of human consciousness and intentionality in the natural world by presenting a (naturalistic) theory of their emergence and gradual evolution by natural selection. In his efforts to integrate them into the scientific worldview, Dennett has openly committed himself to Darwinian adaptationism, granting neither phenomenal consciousness nor intrinsic intentionality a role in the evolution of human cognition (see Zawidzki, 2007, p. 81–85, 122, 126–129; Thompson, 2009, p. 4, 6, 70, 74). Still, this leaves the status of the natural-scientific worldview, its reality and objectivity, unclarified and offers little justification for adopting it. Thompson has suggested that a more comprehensive application of *epoché* and Husserl's notion of constitution could push Dennett to make his "scientific realism" consistent with his analysis of consciousness and theory of stances, for which the reality of objects seems to depend on the adopted stance (Thompson, 2000, 2009, p. 98, 103–105, 152–154).

[27]For these two lines of interpretation of *epoché* and their problems, see Zahavi, 2019.

an answer to both Brentano's position and Dennett's charge of methodological solipsism[28].

The mentalistic aspect of methodological solipsism is encapsulated in Dennett's claim that *epoché* brings "the essence of the mental" to view by "bracketing the real world" (Dennett, 1987, p. 161). Dennett does not elaborate on what he thinks is revealed by this, but one can safely assume that intentionality as an essential structure of conscious experiences is amongst those features. The idea of intentionality as an intrinsic and exclusive feature of *mental* phenomena has its origins in Franz Brentano's work. This is acknowledged by Dennett (1987, p. 67) who speaks of Brentano's Thesis as characterizing intentionality as "the mark of the mental" and the defining feature of mental phenomena: "all and only mental phenomena exhibit intentionality" (Dennett and Haugeland, 1987)[29]. The mentalistic side of methodological solipsism blends with internalism in Dennett's discussion of intentional *objects*. Dennett (1987, p. 153, 161) holds that "the real referents" of our beliefs are more or less inaccessible to introspective autophenomenology of Husserl and Brentano. If shared reality and its real objects are effectively pushed outside the scope of reflective investigations, how can phenomenology account for what is *intended* in conscious experiences? Here, Dennett introduces the fictional element to his reading of methodological solipsism. He proposes that one can make sense of what subject's experiences are *about* by positing a *fictional* world consisting of whatever objects, events, etc. the subject happens to believe in. This notional world can then be described by taking those items as the notional referents of subject's internal representations (Dennett, 1987, p. 118, 153, 155, 158–159, 1991, p. 81). It is in this move from internalism to fictonalism where Brentano's influence is at its strongest. Dennett (1987, p. 153) more or less identifies what he calls notional objects with Brentano's intentional objects. Furthermore, he alludes to Brentano's account of intentional inexistence by acknowledging his merits in discussing the "relation*like* status" of mental phenomena with *non-existent* intentional objects (Dennett, 1987, p. 159) and guaranteeing an odd sort of existence to intentional objects as logical constructs (Dennett, 1978, p. 181).

Challenging the mentalistic, internalist, and representationalist readings of Husserl's account of intentionality (e.g., McIntyre, 1988) in general is beyond the scope of this article[30]. Suffice to say that already in his early philosophy culminating in *Logische Untersuchungen*, Husserl was critical of similar features in Brentano's account[31]. First, Brentano's claim that an intentional relation demarcates psychical phenomena from physical phenomena can be challenged as mentalistic. Not only does it presuppose that all experiences considered "psychical" would take the form of an act directed at an immanent or mental object, but it also excludes phenomena like perceptual sensations (*Empfindungen*) from the mental domain as something "physical" (or, perhaps, physiological) (Husserl, 1984, p. 377–383, 406–407). Second, and relatedly, Brentano's representationalism (i.e., the claim that all psychical phenomena are either representations or founded upon representations) leads him to overlook the non-intentional aspects of experience, such as sensory feelings (e.g., pain and pleasure) and bodily sensations (tactual, visual, olfactory, etc.) (Husserl, 1984, p. 382–383, 406–410). Classifying and determining phenomena in this way delimits conscious experience as a whole, restricting it to one of its subclasses, namely representational psychical acts directed at objects. Thus, Brentano ends up circumscribing the field of descriptive psychology more narrowly than Husserl delineates phenomenological investigations. Husserl's terminological remarks regarding Brentano's expressions pose a further challenge to mentalistic and representationalist readings of intentionality. In short, in Husserl's mind one should avoid speaking of "mental objects," "immanent objectivities" and "mental inexistence," as if the objects intended were intramental parts of experience, enclosed or contained within consciousness (Husserl, 1984, 383–388). Third, Brentano's internalism rests upon an epistemologically and metaphysically unfounded distinction between inner perception and outer perception. This dichotomy leads Brentano to grant inner perception the kind of immediacy and infallibility Dennett ascribes to autophenomenology and to prioritize it over fallible outer perception. Furthermore, the distinction is not metaphysically neutral, since it presupposes an outside world standing against an inner domain of consciousness. By the same token, Husserl claims that psychologists often draw a false antithesis between introspection and outer perception (see Husserl, 1984, p. 13).

More importantly, Dennett does not acknowledge that later in his career Husserl criticized Brentano precisely for the lack of what Husserl calls *intentional analysis* (see Husserl, 1950, p. 86, 1952, p. 59, 1954, p. 237, 1974, p. 252). Husserl maintained that intentional experiences can and should be described reflectively in terms of their non-independent structural features or moments and their correlative intentional objects. For Husserl, intentionality is not mere *aboutness*, as Dennett and others define it (Dennett and Haugeland, 1987). It is not merely a formal feature of directedness of consciousness, namely the fact *that* conscious experience is consciousness *of* something. On the contrary, intentional relation in the pregnant sense includes also *how* and *of what* we are conscious of in different kinds of experience (Husserl, 1976, p. 74; cf. Husserl, 1950, p. 71–72).

In *Ideen I*, Husserl (1976, p. 349) promoted reflecting conscious experience in terms of its hyletic, noetic, and noematic elements. On the one hand, phenomenological reflection may

[28]It is outside the scope of this article to determine whether Dennett's passing remarks on Brentano are any more faithful to the original than his reading of Husserl. I will not take a stand on the accuracy of Husserl's reading of Brentano either.

[29]In contrast to Brentano (and Husserl), Dennett's intentional systems theory decouples intentionality from phenomenal consciousness. In attributing intentional states such as beliefs to others, the intentional stance does not distinguish conscious from unconscious states (Dennett, 2018). Any subject of such intentional attributions whose behavior is reliably predictable from the intentional stance is an *intentional system*, whether it is a human being or other animal, a zombie or a Martian, a corporation or a chess-playing computer (Dennett, 1987, p. 15, 22–23, 28, 58, 1991, p. 76–78).

[30]For a recent discussion, see Drummond (2012), Zahavi (2017, p. 10, 52–53, 79–94, 120).

[31]In the following exposition, I draw from Moran (2000), Cobb-Stevens (2003), Fisette (2010), and Heffernan (2015).

focus on the subjective side of experience by engaging in *noetic* analyses of different modes or types of consciousness (such as perception, imagination, and recollection) that give form and sense to experiences and their modal differences (such as clarity and distinctness). The subjective side also includes the so-called *hyletic* components providing material and content to experiences: the aforementioned perceptual sensations and sensory feelings, but also other types of sensual and bodily experiences related to affect, volition, desire, and sensibility, such as impulses, urges and drives, and even aesthetic pleasure. As non-independent moments of experience, they are not in themselves intentional but they function as "bearers" or "carriers" (*Träger*) of intentionality, and, as such, they partake in establishing and sustaining "consciousness of something" (Husserl, 1976, p. 74–75, 192–194, 1984, p. 409–410). On the other hand, one may attend to the objective side of experience in *noematic* analyses. These investigations focus on how intentional objects in different categories (such as a perceived tree, an imagined centaur, and a remembered event) are given in various modalities of experience (including modalities of being, such as actual and potential, possible and certain, real and fictional, and temporal modalities such as past, present, and future). Intentional analysis, thus, lends itself to two-sided *correlational analysis* between subjective modes of consciousness and intentional objects as their correlates (Husserl, 1950, p. 74–75, 1976, §97).

On the objective side of reflective investigation belong not only the singular objects and states of affairs as they are actually attended; it also entails their background as something implicitly, indeterminately, and potentially co-intended. Husserl extended his noematic analyses to (un)cover these background features of experience or *horizons* as he prefers to call them (see e.g., Husserl, 1950, §20, 1976, p. 188–189). What makes it blatantly clear that he does not view phenomenological reflection as restricted to any kind of mental or internal dimension is that Husserl thinks that the world is the ultimate horizon for every single experience. For Husserl, the world is not merely the spatio-temporal background of every perception and other object-directed intentional acts. It is also the practical and social context of everyday activities imbued with familiarity and historical meaning (Husserl, 1939/1972, p. 49, 52–54, 1950, p. 75, 1954, p. 267). This is why Husserl analyzed the subject of experience as a person embedded in and in relation to its temporal, social, and worldly horizons (see Belt, 2019). This is in stark contrast with the narrow-psychological investigation of subjects isolated from the historical, practical, and social world promoted by methodological solipsism.

Rather than limiting phenomenology to intramental contents, intentional analysis widens the scope of reflective investigations. As Husserl stresses, in phenomenological reflection, understood as intentional analysis, "all occurrences of the life turned *toward the world* [. . .] become accessible to description" (Husserl, 1950, p. 73–74, emphasis added). In intentional analysis Husserl puts into effect, through reflective practice, the methodological change brought about by *epoché*. To repeat, this move should not be understood as turning away from the world and leading back to some kind of internal domain, as commentators interpreting Husserl as a methodological solipsist often claim (see McIntyre, 1988, p. 58–59; Varela et al., 1991/2016, p. 16). Neither should it be framed as *constructing* a subjective world from within nor as *postulating* a sort of fictional parallel world by an external observer, as Dennett (1987, p. 154) would have it. Rather than losing the world, Husserl argues, intentional analysis treats the world as the correlate of all possible conscious experience (Husserl, 1939/1972, p. 46, 1950, p. 75, 1954, p. 235–236). From the perspective of transcendental reflection, intentional analysis is executed as a systematic *constitutional analysis* of all the actual and possible ways of being conscious of objective unities in different categories and regions of objects. This delineation opens up constitutional investigations of different levels of objectivity from psychophysical nature to (inter)personal human reality, regions of value and practical objects, cultural formations like "*state, law, morals, and the church*" (Husserl, 1976, p. 354), ontologies of all the different sciences, and ultimately the whole of what can be called the objective world (Husserl, 1950, p. 85–86, 89, 98, 1976, §80, §86, §149, §§152–153). These phenomena are not only intersubjectively accessible, but also non-private in the more elementary sense of being intersubjectively constituted, i.e., an intentional accomplishment based on and mediated by social interaction (rooted in empathy in particular), communalization, and, in many cases, historical formation over generations[32].

This brief description has hopefully shown that Dennett's interpretation of Husserl as committed to methodological solipsism runs counter to both how Husserl himself views his method and how reflection as intentional analysis is executed. In other words, Husserl is committed to methodological solipsism neither as a research *strategy* nor as an *implemented* method. Consequently, Dennett misses the positive potential of intentional analysis for alleviating the empirical worries raised in the previous sections, most notably underdescription (our tendency to overlook certain features of conscious experience). Let us end this section by returning to Dennett's demonstrations of the allegedly undetected defects in our peripheral vision. Interestingly, Husserl elucidates the indeterminacy and vagueness of experience precisely with concrete examples of visual perception in a series of intentional analyses.

Consider looking at a sheet of paper on the desk in front of you. When you are turned toward the paper, you also perceive objects surrounding it: books, pencils, a cup, etc. As long as you stay focused on the paper, however, it is picked out from its perceptual background. While the surrounding objects also appear as something co-given, they are not perceived attentively but only seen in the background with relative degrees of clarity and saliency. Still, you are at any time free to turn your gaze (*Blick*) from the currently regarded object (e.g., the sheet) to any of those background objects in your visual field (e.g., the teacup on the table) and notice it in passing or concentrate upon it. In shifting focus, you become explicitly conscious of another object and its distinctive features, which were only implicitly seen as potential objects of attention. But at the same time the original act of perception directed at the paper loses its actuality

[32]These lines of investigation are already predelineated in such "egological" introductory works as *Ideen I* (Husserl, 1976, §§151–152) and *Cartesianische Meditationen* (in the Fifth Meditation, Husserl, 1950, §43, §49, §58).

and intuitive clarity and the previously attended object recedes into background (Husserl, 1976, §35, §§44–45, §83). Even such an elementary analysis shows that Husserl did not consider the visual field uniformly detailed or fully determined. Rather, what is actually seen is always accompanied or surrounded by what Husserl, in different passages in *Ideen I*, calls *horizons* of "background inattentiveness" (Husserl, 1976, p. 185–186) and "more or less vague indeterminacy" (Husserl, 1976, p. 91), and "'a halo' of non-actualities" (Husserl, 1976, p. 73).

Now consider Husserl's other well-known example, looking at a tree in a garden through a window (Husserl, 1976, §97, cf. §41). While continuously observing the tree, its manners of appearance may change in various ways: the tree itself might sway in the wind or you can alter your own position in relation to it by tilting your head or taking a step closer to the window to get a better view; you can keep your focus fixed on the tree or let your eyes wander to its branches and trunk; finally, the color of the tree is displayed in a wide range of shades in the changing light. Still, throughout these changes, you observe one and the same tree with prevailing and identical features (such as shape but also color). How do we make sense of this phenomenologically? The first step of intentional analysis is to distinguish how the hyletic moments (e.g., kinesthetic sensations, sensations of color), noetic acts (e.g., perceiving, focusing), and noematic aspects (e.g., the tree as it is seen, the perceived color of its trunk) are correlated in visual experience. Then, constitutional analysis sets out to explicate how a continuous and unified consciousness of an unchanging intentional object or "a synthesis of identification" in Husserl's words is formed in the course of constantly changing perception (Husserl, 1954, p. 160–161, 164).

Husserl's constitutional analysis focuses, first, on the temporality of visual experience. What is currently seen retains the previous phases of perception and anticipates its future course. Only in this way can I see that the momentary changing perspectives are aspects *of this* tree or shades *of this* color; otherwise I would not see the *same* tree if I closed my eyes for a second or be surprised if the tree turns out be a prop when seen from another angle. Husserl also underlines the embodied aspects of seeing and perceiving in general. One's lived body is involved in visual perception both as a center of orientation (the tree is perceived in particular direction, near or far, over there in relation to here, etc.) and as a *locus* of action. Visual experience is not only supported by more or less automatic processes (e.g., my eyes saccade and accommodate distances in tracking the swaying branches), but it also entails an implicit or more explicit sense of motility and awareness of possible lines of voluntary action – a consciousness of "I move" and "I can" in Husserl's terms. These dynamic aspects of perception point to yet another kind of indeterminacy in visual experience and perception in general. Since a tree is always seen from a certain spatial perspective and, thus, presented one-sidedly, it is always open to *further* determinations (there are currently unseen aspects, unnoticed details, etc.) and, correlatively, potential courses of perceptual action: I can *deliberately* change focus or *let* my eyes wander aimlessly, *freely move* around the tree, or force myself to *stay still* and *fix* my eyes more closely on a currently visible detail and so on (Husserl, 1950, §§17–19, 1954, p. 160–161, 164, 1976, §41, §44).

Classical phenomenologists like Husserl and Merleau-Ponty have treated the indeterminacy of vision as an essential feature of perception analyzable in greater detail, not as a contingent cognitive defect. Drawing from these sources, Thompson et al. (1999) demonstrate this by dissecting another of Dennett's examples, namely the thought experiment where one imagines entering a room covered with a wallpaper made out of identical pictures of Marilyn Monroe. They point out several ways in which Dennett mis- and underdescribes the phenomenology of visual perception (see also Printz, 2018; cf. Dennett, 1991, p. 354–355, 359–360).

First, we do not *seem* to see, as Dennett falsely claims, hundreds of Marilyns equally well and in detail. On the contrary, the pictures currently in front of you are seen clearly while posters get less distinct and eventually indistinguishable further to the sides. This is not something undetected at the personal level, as careful phenomenological descriptions offered in terms of horizons and figure-ground structure show. Second, Dennett's depiction of the perceptual situation is artificially static and passive, considering the set-up. In stepping into the room, one has already scanned the environment and, as an active perceiver, is always free to explore it in further detail. Thus, one perceives the room covered with Marilyns, not because each poster presents itself as seemingly distinct at the moment but because the pictures not currently focused upon are still co-present as previously seen copies (that have become indistinct in turning one's head) or potential objects in the future course of perception (that are anticipated but not yet determined). Third, Dennett overlooks the embodied character of visual experience and its spatial configuration: all perceivable things are situated in relation to one's body and its spatial vantage point. What we are facing and where our gaze is focused on partly determine which Marilyns stand out from the background. Moreover, the visual surroundings can be further divided into the immediate spatial context (the more distinct pictures close to the central figure) and the periphery with indeterminate boundaries (it is not easy to tell which posters fall outside of the visual field). In line with Husserl, Thompson et al. (1999) emphasize that the indeterminate background also involves a tacit awareness of one's body (and bodily abilities, I might add, epitomized in the "I can"); every perceivable object is situated in what they, following Merleau-Ponty, call "an implicit bodily space." Along these lines, one could argue that in order to experience the room *covered* with Marilyns, even the surroundings outside of the visual field must be marginally present in perception as something implicitly there. After all, it would surprise us if we turned around to inspect the wall behind our back and the wallpaper looked totally different.

The lessons learned also apply to Dennett's experiment with playing cards, discussed above. When you are asked to stand still and stare straight ahead at a fixed target, it is certainly easier to isolate certain features of central and peripheral vision (or foveal and parafoveal vision) and notice their limitations. Such an illustration may well reveal everyday misconceptions about visual experience and even point out a common blind spot of casual reflection. But it ought to be taken as a prompt for further phenomenological reflection rather than as an indication of unsurpassable errors of reflection, let alone as proof that such phenomena are indescribable from the first-person

perspective. The intentional analyses sketched above can shed light on what is ordinarily overlooked, but they also suggest that Dennett's examples of peripheral vision pay insufficient attention to temporal, embodied, and (en)active aspects of visual *experience*. They illustrate how Husserlian distinctions and lines of investigation can lead to a more comprehensive and faithful description of previously underdescribed phenomena.

Eidetic Variation

Dennett's remaining methodological arguments for phenomenological skepticism all revolve around the question of how phenomenological reflection is related to a single reflecting subject. Arguments stemming from the documented problems of introspective psychology and the perceived deficiencies of introspection in general implicitly assume that Husserlian methodology relies heavily on introspection, i.e., self-observation and the reporting of one's own current or recently past experiences. Along the same lines, Dennett maintains that autophenomenologists treat *oneself* as the sole subject and only object of investigations (the lone-wolf approach) and base their insights exclusively on what can be learned reflectively from one's *own* experience (the first-person-plural presumption). This section aims to undermine these assumptions by setting straight the relationship between phenomenological reflection, the reflecting individual, and her particular experience. This is established by taking a closer look at the eidetic nature of Husserl's reflective phenomenology, especially the methodological procedure called *eidetic variation*. Finally, I will revisit two skepticism-inducing topics of the previous sections, namely cataloging experiences and variation in self-reports.

For Dennett, the perceived inability of introspective psychology to meet the standards of objective science points to a deeper issue inherent in all first-person methods. Since introspection and, by extension, autophenomenology rely on *private* inspection of *one's own particular* experiences, the argument goes, one can never hope to produce reliable results that can be compared, validated, and replicated intersubjectively using first-person methods. Dennett also seems to agree with William James in thinking that introspection always involves *retro*spection. This poses another problem for reflective methodology. As accessing and self-reporting experiences takes time and proceeds in stages, there is always a logical chance of error due to misremembering, no matter how short the time lapse between experiencing and describing it (Dennett, 1991, p. 317–318; cf. James, 1890/1950, p. 185, 189–192).

Husserl was by no means unfamiliar with such skeptical arguments. A manuscript drafted on Husserl's behalf by his assistant Edith Stein to address their contemporary psychologists, Theodor Elsenhans and August Messer, even seems to endorse a similar train of thought:

> "What is genuinely psychic [. . .] cannot be treated in the same way as external objects. A perception, a feeling of joy, a simple sensation, flows away; and once it has decayed, it has *irretrievably* [*unwiederbringlich*] disappeared. I cannot hold on to them and inspect [*vorzeigen*] them, so as to give some determinacy to the fluid descriptive concepts corresponding to them; I cannot hold them up to each other, so as to isolate common attributes and, with their help, to form classificational concepts. [. . .] I have a flux of unrepeatable [*unwiederholbaren*] and incomparable [*unvergleichbaren*] individualities, which mock any kind of conceptual grasp. *A pure empirical science* [*Erfahrungswissenschaft*] *of the psychic is utterly impossible.*"[33]

Husserl scrutinized the limits of attaining reflective knowledge about the particularities of consciousness also in his published writings. Due to their flowing and fluctuating character, he argues in *Ideen I*, individual experiences can never be completely perceived and fully grasped in reflection; nor can I inspect *my* stream of consciousness in its entirety in the present moment by "swimming after" it retrospectively (Husserl, 1976, p. 93–94, 96, 156–157). How can phenomenological reflection and phenomenology as a discipline, then, claim to overcome these problems?

The basic idea of eidetic phenomenology is simple enough. From an empirical-psychological standpoint, Husserl (1984, p. 6–7, 12–13) maintains, experiences are perceived and treated as particular facts, classes of real events, mental or psychological attributes or the like. Husserl proposes a methodological re-orientation or a change of perspective called *eidetic reduction*, which leads phenomenology to consider experiences according to their essential features and necessary connections (pure essences or *eide* in Husserl's terminology) (Husserl, 1976, p. 6, 8). Instead of trying to document currently ongoing or previously had individual experiences in an attempt to establish inductive generalizations and empirical classifications, phenomenological reflection sets out to uncover, intuitively apprehend, and analyze the essences of different kinds of experience and their essential relations on different levels of generality and specificity. That is, eidetic investigations focus not only on experience in general as the highest genus, but also on perception, remembering, willing, empathy, etc. as its subordinate kinds or species (Husserl, 1976, p. 30, 157). In short, Husserl presents an alternative to introspective and experimental psychology of his time by developing phenomenology as an eidetic science or a "science of essences" (*Wesenswissenschaft*).

How does eidetic phenomenology examine its subject matter? Husserl (1976, p. 13, 15–16, 69) believes that essences can be intuitively exemplified with both actual and possible instances, no matter if they are currently perceived, remembered or "merely" imagined. Conversely, one is always free to shift focus from a (real or imagined) particular experience or an experience of something singular to corresponding essences in an act called *ideation*. Husserl often illustrates his eidetic method with simple exemplary analyses of perceptual phenomena (such as perceiving a table, hearing a sound, or seeing colors), but ideation is in principle applicable to all kinds of objects of experience from spatial shapes to social processes (Husserl,

[33] Husserl (1987, p. 234), Husserl and Stein (2018, p. 455), translation modified; cf. Husserl (1954, p. 181). Reflection being "ever retrospective" was one of the root causes of skeptical concerns raised by yet another contemporary psychologist, Henry J. Watt, whose arguments Husserl (1976, §79) scrutinizes in *Ideen I*. For the parallels between Watt's and Dennett's skepticism, see Cerbone (2003, p. 124–128, 2012, p. 12–16). For Husserl's response, see also Cai (2011, p. 51–54, 90–92), Zahavi (2015).

1976, §4; see also Husserl, 1950, §34, 1962, §9a,e, §10). Husserl further assures that ideation (or "seeing essences") is nothing mysterious or metaphysically compromising; rather, it is familiar to anyone who has learned mathematics and actively acquired first-hand geometrical insights (in intuitively determining, for instance, that circle is a type of conic section) (Husserl, 1962, p. 85, 87, 1976, p. 47–48). Pushing the analogy further, he holds that eidetic phenomenology demands working freely in imagination like a geometer who produces figures in phantasy, reshapes them at will, and runs through any number of conceivable configurations to probe geometrical insights (Husserl, 1976, §4, §7, §70). To preserve the freedom and presuppositionlessness of eidetic research, ideation strives to discern essences *purely*, that is, without positing the actuality of their corresponding particulars. The universality of eidetic claims is, therefore, not restricted to actual cases or to what is factually possible (as is the case with empirical generalizations); they extend to all *conceivable* experiences or "pure possibilities" in Husserl's vocabulary (Husserl, 1939/1972, §82, §86, §§89–90). How is such a transition from particular experiences and empirical generalities to intuitively apprehended pure essences and essential structures supposed to be accomplished reflectively? What is the explicit methodic form of ideation?

As an eidetic method, ideation can be articulated as a procedure called *eidetic variation*[34]. The process involves using imagination and it proceeds in stages. One starts with an actual, remembered, or imagined example considered as an instantiation of a certain type of experience. Taking it as a guiding model, one then modifies its features freely and as far as possible in imagination in order to produce an open-ended series of *variants* of the same type. Running through all the different variations, one is finally supposed to be able to discern and single out their overlapping or coinciding features and to obtain an intuitive grasp of what stays *invariant* throughout the series. This is what Husserl calls *eidos*, pure or universal essence, and necessary or universal form – in short, it is something without which the experience in question is inconceivable. After grasping or "seeing" essences in this way, the resulting eidetic findings can then be conceptualized, further analyzed, and expressed in the form of universal statements or "eidetic laws" in Husserl's terms.

The previous subsection already provided eidetic descriptions at the highest level of generality. Intentionality was presented as an essential feature of conscious experience in general and further divided into its noetic, hyletic, and noematic moments and horizon structures. These basic distinctions were then applied, by way of example, to concrete cases in intentional analyses focused on the indeterminacy of visual experience. Another essential feature of consciousness touched upon is its temporal structure. Not only are intentional objects of experience constituted as something "fixed and abiding" over time and through changing experiences, but also consciousness itself is temporally constituted (see Husserl, 1950, §20). A melody, for instance, is a temporal object with duration, but our conscious experience of it endures as well. To hear an array of sounds *as* a melody requires that past tones are somehow retained (as just passed) and a succession of chords is implicitly anticipated (as soon to arrive) in what is heard in the current moment; a certain note can appear as a discord only in contrast to such a concordant continuum. Husserl considered it an essential feature of consciousness that each phase of experience has a threefold structure of primal impression, retention, and protention, which unifies conscious experience passing from one now-moment to another and gives a flowing character to it. By virtue of such temporal form, individual episodes have duration (e.g., a feeling of joy initially rises, intensifies, and dissipates in phases, and eventually fades into past), but it also binds conscious experiences together, and this connection of experiences is temporally ordered into successive and simultaneous experiences (Husserl, 1966, p. 66, 72–73, 87, 313–317, 323–324, 1976, §§81–82). This is why Husserl can ultimately claim that "*conscious life as a whole* [. . .] *is synthetically unified*" (Husserl, 1950, p. 80) and that each experience belongs to a "*single* endless '*stream of experience*'" (Husserl, 1976, p. 182).

The topic of inner time-consciousness as a basic structure of experience and the fundamental form of synthesis led Husserl to notoriously difficult in-depth explorations whose technical intricacies need not concern us here. The fact that Husserl investigated intentionality and temporality as universal *structures* of conscious experience is already sufficient to demonstrate that eidetic descriptions are by no means restricted to *eide* of different categories and regions of *objects* and their essential features. Correlational and constitutional analyses also strive to explicate eidetically how the subjective and objective sides of experience are necessarily connected and how experiences are *inter*connected in a structured, lawlike manner.

As an eidetic method, phenomenological reflection circumvents the perceived main weaknesses of introspection. First, eidetic phenomenological description is by no means restricted to what is experienced here and now. Neither does it rely on trying to faithfully retrieve already passed and irrevocably faded individual experiences. Since the same essential and structural features are instantiated by countless examples, one can replicate the process of eidetic variation by finding another starting exemplar, producing a new series of variants and (re)evoking eidetic insights[35]. Second, the results are not incommensurable, because one is not reporting individual or private events. Rather, eidetic descriptions are concerned with shared structures of experience and they claim universal validity. They can, therefore, always be compared to and challenged by competing descriptions in terms of clarity, accuracy, scope, amount of detail, ability to differentiate, etc. Third, in arriving at eidetic claims, phenomenological reflection does not simply draw from one's own experience and (over)generalize. On the contrary, confining eidetic variation to what the reflecting individual has experienced first-hand would seriously constrain our ability to probe what is conceivable. Since our imaginative abilities are also limited, Husserl advocates "pollinating" imagination with

[34]This reconstruction of eidetic variation is based on Husserl (1939/1972, §87a,e, 1962, §9a,c,e, 1974, §98, 1987, p. 245).

[35]Husserl (1974, p. 255, 1976, §75, cf. §7, §34, 1987, p. 234–235, 245). Cerbone (2003, 2012) offers a more detailed answer to Dennett's argument concerning error from misremembering.

experiential, historical, and even fictional sources depicting lived experience before engaging in eidetic investigations (see Husserl, 1952, p. 51–53, 1976, §70). Enriching our imagination with an array of intuitive material to work with extends the scope of variation to unthought-of but still conceivable possibilities and illuminates unnoticed or underdescribed features of conscious experience. One could even argue that this addresses one of Dennett's recurrent worries, namely that philosophers have a tendency of "mistaking a failure of imagination for an insight into necessity" (Dennett, 1991, p. 401, cf. 48, 440).

Dennett's conception of autophenomenology fails to recognize the eidetic aspect of phenomenological reflection. This omission is echoed by what is missing from his heterophenomenological alternative. In Dennett's reading, the goal of autophenomenology is to characterize one's *own* notional world introspectively "from the inside." By the same token, heterophenomenology sets out to describe the notional world of *another* subject "from the outside" by interviewing and observing them. In both cases, one starts by extracting a *single* person's account of their own experience (see Dennett, 1987, p. 153, 158; cf. Dennett, 2003, 2007). But how do we generalize from such accounts? Dennett accuses autophenomenologists of carelessly extrapolating from their own experience and simply assuming that the same first-person descriptions are reproducible by others. In heterophenomenology, by contrast, test subjects are deliberately steered away from theorization and "faux generalization" (see Hurlburt and Schwitzgebel, 2007, p. 127, 255) when giving their subjective accounts. It follows that individual reports are only subsequently compared, interpreted, and cataloged. Yet, Dennett says little about how one extrapolates *post facto* from a set of heterophenomenological texts. He is fast to deny that first-person data is "averaged out" by statistical means (Dennett, 2007), but how should one draw conclusions if and when variation between subjects ensues? Dennett (2003, 2007) rightly emphasizes the role of *interpretation* already present in identifying first-person reports, turning them into useful data, and finally using them as a source of evidence. But his remarks on the requirements of interpretation are ambiguous at best. What background knowledge does a researcher need, and is entitled to use, in interpretation? Do heterophenomenologists apply the same vocabulary as test subjects or adopt another terminological or theoretical framework for interpretation? How are the findings ultimately classified?

These are pressing questions, not least because heterophenomenology is supposed to provide a neutral inventory of phenomenological items or a "heterophenomenological catalog" – something Husserl's phenomenology is supposedly unable to deliver. Dennett (1991, p. 45–46) is the first to admit that his own provisional classification of inner, outer, and affective experiences is based on "dubious tradition" and "superficial similarities" rather than a "deep kinship" between phenomena. Presumably he favors heterophenomenological reports that can be interpreted indirectly and antecedently so that one can refrain from committing to pre-established categories. However, it is unclear how the heterophenomenological approach could avoid such pitfalls since interpretation of test-subjects' testimonies relies on the intentional stance of the observer. It is hard to see how identifying, describing, and classifying the contents of other people's reports is possible without resorting to what the interpreter has learned pre-reflectively from what she has lived through and what she already reflectively knows about her own experience (Gallagher, 1997; Carr, 1998; Marbach, 2007). Following James (1890/1950, p. 194–196), one could also argue that describing consciousness is particularly vulnerable to linguistic influences, since we are prone to use "the vocabulary of outward things" and to suppose substantive entities; we also tend to overlook and misconstrue conscious phenomena due to lack of words and "the dependence of psychology on common speech" (cf. Husserl, 1984, p. 15). Consequently, the less informed one is about phenomenologically attuned and reflectively secured distinctions, the more interpretation is guided by preconceived conceptual categories, associative typifications, folk-psychology, and other potential sources of bias. If some kind of taxonomy of conscious experiences is needed for resolving reflective disagreements, as Dennett insists, Husserlian phenomenology offers an invaluable source.

On this point Dennett has partly conceded to criticism. He admits having previously ignored "data" acquired by reflecting on structures of consciousness from the first-person perspective. Moreover, he is happy to conclude that subtle phenomenological distinctions Husserl, among others, provided can be put to good use in conducting heterophenomenological interviews. Phenomenology can enrich the vocabulary and "tease out" aspects of experience at the personal level (Dennett, 2007). This is a step in the right direction. For all intents and purposes, Dennett here acknowledges that Husserlian phenomenology helps to tackle under- and misdescription of conscious experience with terminological and analytical tools.

For the skeptic, interpersonal and intrapersonal variation in introspective self-reports is indicative of the unreliability of first-person methods (see section "Dennett's Empirical Arguments"). How does phenomenological reflection fare against the argument from variation? One option is to argue that eidetic phenomenology is not vulnerable to whatever sources of error variation in psychological self-observations might indicate by insisting on the differences in establishing and validating eidetic and empirical claims. Simply put, phenomenological reflection neither relies on cumulative results of personal introspection nor bases its claims on inductive reasoning or statistical inference using data collected from untutored test subjects (or surveying the researchers for that matter). Husserl maintains that since eidetic variation operates freely in imagination, without presupposing or positing the actuality of its examples, it should not be mistaken for "empirical variation" restricted to and constrained by factual cases, let alone required to seek experiential confirmation for its factual basis (see Husserl, 1952, p. 47–48, 51, 54, 1974, p. 218–219, 255, 1976, p. 171–172). It would seem to follow that potential distortions in gathering "first-person data" do not carry over to eidetic phenomenology and observed variation (whatever its ultimate cause) has no bearing on phenomenological reflection.

This line of counterargument, however, oversimplifies the relationship between eidetic and empirical knowledge and misses the potential usefulness of (f)actual variation for eidetic

phenomenology. Husserl (1939/1972, p. 423, 426, 1962, p. 71, 74, 86) states clearly that every actual occurrence can be turned into a variant and considered as an example, treating it as one pure possibility among others. It follows that eidetic phenomenology can accommodate empirical material by incorporating it into eidetic variation as starting examples or potential variants. In principle, it makes no difference whether the presumed variation is exposed by first-person, second-person, or third-person investigations, as long as the findings are transformed into intuitively imaginable *possible* experiences. This already shows that neither observed variation nor empirical findings in general should be outrightly dismissed or ignored as irrelevant to eidetic claims. More to the point, eidetic phenomenology can accommodate cases where there *are* good reasons to believe that observed variation points to real underlying differences in how we experience things. Zahavi, among others, has suggested that especially the person-level descriptions of real-life deviations and anomalous cases studied in fields such as psychopathology, cognitive and developmental psychology, neurology and anthropology can both challenge our universalistic eidetic claims (as potential empirical counterexamples) and provide illuminating cases for modifying them. Exceptional cases and human variation in general may, then, prompt phenomenologists to revise and refine eidetic descriptions by "pollinating" imagination and extending the scope of eidetic variation as outlined above (see Zahavi, 2017, p. 151–156).

Intersubjective Validation

In section "Dennett's Methodological Arguments," Dennett's motivation for methodological skepticism was crystallized in the three requirements of scientificity that first-person methods supposedly fail to meet: publicity or intersubjectivity, reliability, and agreement. Above, I have argued that phenomenological reflection is neither a solipsistic nor an introspective technique. This should clear away the main obstacle for thinking that Husserlian phenomenology cannot meet the standards of publicity or intersubjectivity. Simply put, Husserl's reflective methodology does not investigate conscious experiences as intersubjectively inaccessible phenomena. On the contrary, the interpretation of *epoché* and intentional analysis defended above shows that the scope of phenomenological reflection is extended to the shared world with its publicly available and intersubjectively constituted objects (rather than restricted to any inner, mental, or private domain). Furthermore, eidetic reflection aims to discover essential or structural features of experience (not facts about any single consciousness or private events). Commentators have also stressed the role of language, shared terminology, and communicative efforts in making phenomenological descriptions public and open to mutual criticism from the get-go (phenomenological descriptions are based neither on private language nor any non-linguistic means) (Sokolowski, 2008; cf. Cai, 2011, p. 126–128, 154–155; see also Zaner, 1973). In this way, the proper domain of phenomenological reflection is, in principle, accessible to everyone. However, in Husserl's view, phenomenological descriptions claim intersubjective validity also in a stronger sense.

For Husserl, the final validity of phenomenological descriptions does not rest on what Dennett calls lone-wolf autophenomenology. The objectivity of phenomenological results is ultimately decided by an intersubjective communal practice, rather than simply *presuming* that others will agree with a subjective account of mine, yours, or anyone else. Husserl (1950, p. 47) states clearly that only the results that can stand the test of mutual clarification and critique can be deemed "objectively valid." This lengthy passage from *Ideen I* captures both the requirements and potential benefits of phenomenological reflection carried out intersubjectively:

> "If one has acquired the right attitude and fortified it through practice, but, above all, if one has gathered the courage to follow the clear instances of essential givenness in a radically unprejudiced manner [*in radikaler Vorurteilslosigkeit*], untroubled by all the currently circulating and learned theories, then firm results quickly ensue, results that are the same for everyone in the same attitude; there arise substantial possibilities of communicating to others what one has seen oneself, testing [*nachprüfen*] their descriptions, bringing out the unnoticed intrusions of empty verbal meanings, and, through subsequent measuring [*Nachmessung*] in intuition, making known and eradicating errors that are possible here as they are in every sphere concerned with validation."[36]

According to Husserl, then, the possibility of reaching shared results is opened by (1) adopting the phenomenological attitude, (2) sufficient training, and (3) freedom from presuppositions (all supported by *epoché*) while accepting only what is (4) intuitively given in reflection. What is equally important is (5) sharing one's findings and (6) testing or verifying other people's results reflectively in order to (7) identify biases and mistakes and to (8) correct errors. While such a general characterization hardly passes as a step-by-step guideline for conducting phenomenological research, it demonstrates that validating results in a scientific community that shares basic methodology, vocabulary, and research practices is quintessential for Husserlian phenomenological reflection. In fact, Husserl was convinced that real advances in reflective phenomenology required generations of researchers all committed to a shared goal, mutual criticism, and taking over others' work (see Husserl, 1974, p. 36, 1984, p. 16–17). Naturally, there is no way to guarantee that controversies can always be settled and contradictions solved, but, with the shared methodological *praxis* outlined in this section, finding common ground for handling the disputes is much more typical than Dennett gives Husserlian phenomenology credit for.

As should be clear by now, Husserl sees phenomenological reflection neither as an infallible nor an incorrigible process. In particular, Husserl does not appeal to immunity from error on the basis of what Roy called the "double thesis," namely that consciousness as a whole is both readily accessible and fully transparent to reflection. This is shown by Husserl's recurring comments about our inability to ever grasp the particularities of flowing and fluctuating conscious experience in its entirety. Husserl voices similar

[36] Husserl (1976, p. 201, 2014, p. 173), translation modified.

reservations, for instance, in scrutinizing the indubitability of reflective knowledge of the self[37]. While eidetic methodology seeks to elevate phenomenological reflection from the individual and particular to the essential and structural features through ideation, the eidetic procedure by no means *secures* infallible results either. As a repeatable and open-ended process that can incorporate both challenging and illuminating material, eidetic variation, rather, invites a continuous refining of phenomenological descriptions. This is closer to an ideal of science as a fallible but self-correcting endeavor than to the kind of commitment to infallibility and incorrigibility Dennett ascribes to post-Cartesian first-person investigations. In fact, in *Logische Untersuchungen*, Husserl (1984, p. 15–17) discusses difficulties involved in stating and communicating results in such a way that once-acquired phenomenological insights can be reidentified, tested, and confirmed by others well-versed in phenomenological methodology. Overcoming such obstacles is a prerequisite for conceiving phenomenology as a *scientific* philosophy.

According to Husserl's guidelines, assessing phenomenological results intersubjectively demands shared methodology and terminology, a certain attitude and training, and a cooperative research community open to mutual criticism. How do such requirements square with Dennett's claim that phenomenology has failed to come up with "a single, settled method" everyone agrees upon? In this section, I have laid out the basic elements of phenomenological reflection to challenge the perception that Husserlian phenomenology is lacking in methodological foundations. While the methodological features are by no means uncontested even within the phenomenological tradition and their details can certainly be challenged, I do not see how *this sort* of methodological debate would merit Dennett's wholesale methodological skepticism about (auto)phenomenology. It rather seems that Dennett's call for methodological *unanimity* turns out to be too strong a requirement. By the same standards, his heterophenomenological alternative would hardly pass as a viable method, as the debates surrounding its nature and general acceptance (see Zahavi, 2007), as well as the above-discussed ambiguities concerning interpretation, generalization and classification, show.

When it comes to the related claim that phenomenological methodology has failed to produce agreement about its results, unanimity presents an equally problematic criterion. As Husserl already argued in response to his contemporary critics, if our reflective insights have to be unanimously affirmed to be considered legitimate, the same standard would render all experiential evidence questionable, since both intuition and *Erfahrung* can and have been appealed to tentatively and even arbitrarily[38]. Revisions occur even in such eidetic disciplines as mathematics and logic, but it hardly undermines the *possibility* of attaining firm results and *ideally* even complete evidence. The pursuit of eidetic knowledge does not imply infallibility nor claim freedom from error in phenomenology either[39]. In *Formale und transzendentale Logik* (1929), Husserl states that the possibility of deception pertains to *every* kind of evidence; even ostensibly apodictic evidence can be annulled by further evidence (Husserl, 1974, p. 164). The point is that it takes more than allusions to disagreements, undecided cases, and occasional errors to establish that phenomenological reflection is an *un*reliable method. Dropping the insistence on full agreement pushes the proponents of phenomenological skepticism to specify what counts as a "sufficient degree" of reliability and objectivity for reflective knowledge and how phenomenological reflection purportedly fails to meet those standards. Anecdotal evidence of "the battle of 'intuitions'" simply won't cut it.

CONCLUSION

In this article, I provided arguments to dispute Dennett's methodological claims that Husserlian phenomenology is committed to introspection, methodological solipsism, the first-person-plural presumption, and the lone-wolf approach. In parallel, I suggested how *epoché*, intentional analysis, eidetic variation, and intersubjective validation serve to alleviate the more empirical worries about overinterpretation, underdescription, and disagreement. I concluded by addressing Dennett's assumption that phenomenological reflection fails to meet at least three criteria of scientificity, namely publicity or intersubjectivity, reliability, and agreement. What is the outcome of these considerations for phenomenological skepticism motivated by the above-mentioned empirical and methodological reasons?

The strong version of phenomenological skepticism is untenable. It pushes Dennett to adopt metaphysical minimalism and even illusionism, rather than securing the coveted ontological neutrality. The *categorical* version of weak skepticism also loses its appeal, since Husserlian phenomenology is not committed to the doctrine of infallibility and its methodology supports corrigibility in practice. There is simply no reason to question the possibility of reflective knowledge *in general* on the grounds that phenomenological reflection can err and its results are open to modifications.

What about the gradual version of weak skepticism? Husserl explicitly acknowledged the elusive nature of conscious

[37]Husserl differentiates between (ongoing) experiences, abilities, and (habitual) dispositions as individual contents of the concrete ego that are all accessible to reflection but with different levels of clarity and certainty. Even though the ego's "living self-presence" (*lebendige Selbstgegenwart*) and its temporal structure might be indubitable, its own past (*Selbstvergangenheit*) is given indeterminately or even obscurely, and its abilities and dispositions are not indubitable regarding their details (Husserl, 1950, p. 61–62, 67). Husserl (1950, p. 62) even alludes to the same kind of skeptical questions concerning self-deception as Dennett: "How far can the transcendental ego [*transzendentale Ich*] be deceived about itself and how far do the absolutely indubitable components extend despite the possible deception?" Knowledge about *what* is indubitable in self-experience and *how far* the clarity of "I am" reaches must be determined critically; knowledge about such questions can only be attained as a "critical achievement" (see Husserl, 2002, p. 401–402; Cai, 2011, p. 110–111).

[38]See Husserl (1976, §79) answering psychologist Theodor Ziehen's skeptical remarks (cf. Cerbone, 2003, p. 128, 132).

[39]See again the manuscript addressing Elsenhans and Messer (Husserl, 1987, p. 246–247).

experience and identified many of the problems associated with introspection and casual reflection that motivate weaker forms of skepticism. But what really confirms Husserl's lack of trust in *untutored* reflection is that he went to such great lengths to hone his methodology in order to safeguard phenomenological reflection against such shortcomings. In light of the methodological considerations provided in this article, it seems premature to conclude, *pace* Roy, that Husserl's phenomenology, in its original form, cannot tolerate a certain degree of fallibility. In recognizing *and* striving to overcome the limitations of reflection in order to attain reflective knowledge, Husserl's methodology is, rather, well-positioned to alleviate the worries expressed by gradual weak skepticism.

AUTHOR CONTRIBUTIONS

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

FUNDING

This work was supported by the Kone Foundation and the Finnish Cultural Foundation.

ACKNOWLEDGMENTS

An earlier version of this article was presented at the CFS Research Seminar in Copenhagen. I would like to thank Dan Zahavi and all the participants for their valuable feedback. I wish to also thank Sara Heinämaa, Susanna Lindberg, and the participants of the Phenomenology Research Seminar in Helsinki for their helpful suggestions on previous drafts. I am especially grateful to Mirja Hartimo, Leila Haaparanta, and Fredrik Westerlund for fruitful discussions and invaluable comments on the manuscript. Finally, I would like to thank both reviewers and the handling editor CG for their insightful comments and Lucy Osler for all the help with language editing.

REFERENCES

Belt, J. (2019). Between minimal self and narrative self: a Husserlian analysis of person. *J. Br. Soc. Phenomenol.* 50, 305–323. doi: 10.1080/00071773.2019.1577067

Cai, W. (2011). *Reflection as a Form of Human Life: Methodological Issues in Phenomenology*. Copenhagen: Faculty of Humanities, University of Copenhagen.

Carr, D. (1998). Phenomenology and fiction in Dennett. *Int. J. Philos. Stud.* 6, 331–344. doi: 10.1080/096725598342028

Cerbone, D. R. (2003). "Phenomenology: straight and hetero," in *A House Divided: Comparing Analytic and Continental Philosophy*, ed. C. G. Prado (Amherst, NY: Humanity Books), 105–138.

Cerbone, D. R. (2012). "Phenomenological method: reflection, introspection, and skepticism," in *The Oxford Handbook of Contemporary Phenomenology*, ed. D. Zahavi (Oxford: Oxford University Press), 7–24.

Cobb-Stevens, R. (2003). "Husserl's theory of judgment: a critique of Brentano and Frege," in *Husserl's Logical Investigations Reconsidered*, ed. D. Fisette (Dordrecht/Boston, MA: Kluwer), 151–162.

Crane, T. (2017). Papineau vs Dennett: a philosophical dispute. *Times Literary Supplement* 5966, 15–16.

Dennett, D. C. (1978). *Brainstorms*, 2nd Edn. Cambridge, MA: MIT Press.

Dennett, D. C. (1987). *The Intentional Stance*. Cambridge, MA: MIT Press.

Dennett, D. C. (1988). "Quining qualia," in *Consciousness in Contemporary Science*, eds A. Marcel and E. Bisiach (Oxford: Oxford University Press), 42–77. doi: 10.1093/acprof:oso/9780198522379.003.0003

Dennett, D. C. (1991). *Consciousness Explained*. New York, NY: Back Bay Books.

Dennett, D. C. (1992). "The self as a center of narrative gravity," in *Self and Consciousness: Multiple Perspectives*, eds F. S. Kessel, et al. (Hilsdale, NJ: Erlbaum), 103–115.

Dennett, D. C. (1993). Caveat emptor. *Conscious. Cogn.* 2, 48–57.

Dennett, D. C. (2002). How could I be wrong, how wrong could I be? *J. Conscious. Stud.* 9, 13–16.

Dennett, D. C. (2003). Who's on first? Heterophenomenology explained. *J. Conscious. Stud.* 10, 19–30.

Dennett, D. C. (2007). Heterophenomenology reconsidered. *Phenomenol. Cogn. Sci.* 6, 247–270. doi: 10.1007/s11097-006-9044-9

Dennett, D. C. (2016). Illusionism as the obvious default theory of consciousness. *J. Conscious. Stud.* 23, 65–72.

Dennett, D. C. (2018). "Reflections on David Rosenthal," in *The Philosophy of Daniel Dennett*, ed. B. Huebner (Oxford: Oxford University Press), 165–170.

Dennett, D. C. (2020). A history of qualia. *Topoi* 39, 5–12. doi: 10.1007/s11245-017-9508-2

Dennett, D. C., and Haugeland, J. (1987). "Intentionality," in *The Oxford Companion to the Mind*, ed. R. L. Gregory (Oxford: Oxford University Press).

Dennett, D. C. (1969/1986). *Content and Consciousness*, 2nd Edn. London & New York, NY: Routledge.

Dennett, D. C. (2001/2018). "The fantasy of first-person science," in *The Map and the Territory: Exploring the Foundations of Science, Thought and Reality*, eds S. Wuppuluri and F. Doria (Cham: Springer), 455–473.

Dreyfus, H. (1980). Dasein's revenge: methodological solipsism as an unsuccessful escape strategy in psychology. *Behav. Brain Sci.* 3, 78–79. doi: 10.1017/s0140525x00001849

Drummond, J. J. (2012). "Intentionality without representationalism," in *The Oxford Handbook of Contemporary Phenomenomenology*, ed. D. Zahavi (Oxford: Oxford University Press), 115–133.

Fisette, D. (2010). "Descriptive psychology and natural sciences: Husserl's early criticism of Brentano," in *Philosophy, Phenomenology, Sciences: Essays in Commemoration of Edmund Husserl*, eds C. Ierna, et al. (Dordrecht: Springer), 221–253. doi: 10.1007/978-94-007-0071-0_10

Fodor, J. (1980). Methodological solipsism considered as a research strategy in cognitive psychology. *Behav. Brain Sci.* 3, 63–73. doi: 10.1017/s0140525x00001771

Frankish, K. (2016). Illusionism as a theory of consciousness. *J. Conscious. Stud.* 23, 11–39.

Gallagher, S. (1997). Mutual enlightenment: recent phenomenology in cognitive science. *J. Conscious. Stud.* 4, 195–214.

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

Goldman, A. (1997). Science, publicity, and consciousness. *Philos. Sci.* 64, 525–545. doi: 10.1086/392570

Gutland, C. (2018). Husserlian phenomenology as a kind of introspection. *Front. Psychol.* 9:896. doi: 10.3389/fpsyg.2018.00896

Heffernan, G. (2015). The Paradox of objectless presentations in early phenomenology: a brief history of the intentional object from Bolzano to Husserl, with concise analyses of the positions of Brentano, Frege, Twardowski, and Meinong. *Stud. Phaenomenol.* 15, 67–91. doi: 10.5840/studphaen2015155

Hurlburt, R. T., and Schwitzgebel, E. (2007). *Describing Inner Experience? Proponent Meets Skeptic*. Cambridge, MA: MIT Press.

Husserl, E. (1950). *Cartesianische Meditationen und Pariser Vorträge*, ed. S. Strasser. Husserliana I. Den. Haag: Martinus Nijhoff.

Husserl, E. (1952). *Ideen zu einer reinen Phänomenologie und phänomenologischen Philosophie. Drittes Buch: Die Phänomenologie und die Fundamente der Wissenschaften*, ed. M. Biemel. Husserliana V. Den Haag: Martinus Nijhoff.

Husserl, E. (1954). *Die Krisis der europäischen Wissenschaften und die transzendentale Phänomenologie*, ed. W. Biemel. Husserliana VI. Den Haag: Martinus Nijhoff.

Husserl, E. (1962). *Phänomenologische Psychologie. Vorlesungen Sommersemester 1925*, ed. W. Biemel. Husserliana IX. Den Haag: Martinus Nijhoff.

Husserl, E. (1966). *Analysen zur passiven Synthesis. Aus Vorlesungs- und Forschungsmanuskripten 1918–1926*, ed. M. Fleischer. Husserliana XI. Den Haag: Martinus Nijhoff.

Husserl, E. (1974). *Formale und transzendentale Logik. Versuch einer Kritik der logischen Vernunft*, ed. P. Janssen. Husserliana XVII. Den Haag: Martinus Nijhoff.

Husserl, E. (1975). *Logische Untersuchungen. Erster Band: Prolegomena zur reinen Logik*, ed. E. Holenstein. Husserliana XVIII. Den Haag: Martinus Nijhoff.

Husserl, E. (1976). *Ideen zu einer reinen Phänomenologie und phänomenologischen Philosophie. Erstes Buch: Allgemeine Einführung in die reine Phänomenologie. 1. Halbband*, ed. K. Schuhmann. Husserliana III/1. Den Haag: Martinus Nijhoff.

Husserl, E. (1984). *Logische Untersuchungen. Zweiter Band, Erster Teil: Untersuchungen zur Phänomenologie und Theorie der Erkenntnis*, ed. U. Panzer. Husserliana XIX/1. Den Haag: Martinus Nijhoff.

Husserl, E. (1987). *Aufsätze und Vorträge (1911–1921)*, eds T. Nenon and H. R. Sepp. Husserliana XXV. Den Haag: Martinus Nijhoff.

Husserl, E. (2002). *Einleitung in die Philosophie. Vorlesungen 1922/23*, ed. B. Goossens. Husserliana XXXV. Dordrecht: Springer.

Husserl, E. (2014). *Ideas for a Pure Phenomenology and Phenomenological Philosophy. First Book: General Introduction to Pure Phenomenology*, trans. D. O. Dahlstrom. Indianapolis, IN: Hackett.

Husserl, E., and Stein, E. (2018). "Critique of theodor elsenhans and august messer (1917) (Edith Stein's draft)," in *The Sources of Husserl's 'Ideas I'*, eds A. Staiti and E. Clarke, trans. E. Clark (Berlin: de Gruyter), 449–468. doi: 10.1515/9783110551594-029

Husserl, E. (1939/1972). *Erfahrung und Urteil. Untersuchungen zur Genealogie der Logik*, ed. L. Landgrebe. Hamburg: Felix Meiner.

James, W. (1890/1950). *The Principles of Psychology*. New York, NY: Dover.

Kotchoubey, B., Tretter, F., Braun, H. A., Buchheim, T., Draguhn, A., Fuchs, T., et al. (2016). Methodological problems on the way to integrative human neuroscience. *Front. Integr. Neurosci.* 10:41. doi: 10.3389/fnint.2016.00041

Lutz, A. (2002). Toward a neurophenomenology as an account of generative passages: a first empirical case study. *Phenomenol. Cogn. Sci.* 1, 133–167. doi: 10.1023/A:1020320221083

Marbach, E. (2007). No heterophenomenology without autophenomenology: variations on a theme of mine. *Phenomenol. Cogn. Sci.* 6, 75–87. doi: 10.1007/s11097-006-9027-x

McIntyre, R. (1988). "Husserl and the representational theory of mind," in *Perspectives on Mind*, eds H. R. Otto and J. A. Tuedio (Dordrecht: D. Reidel), 57–76. doi: 10.1007/978-94-009-4033-8_6

Moran, D. (2000). Husserl's critique of Brentano in the Logical Investigations. *Manuscrito* 23, 163–205.

Nisbett, R. E., and Wilson, T. D. (1977). Telling more than we can know: verbal reports on mental processes. *Psychol. Rev.* 84, 231–259. doi: 10.1037/0033-295x.84.3.231

Overgaard, M., Gallagher, S., and Ramsøy, T. Z. (2008). An integration of first-person methodologies in cognitive sciences. *J. Conscious. Stud.* 15, 100–120.

Peels, R. (2016). The empirical case against introspection. *Philos. Stud.* 173, 2461–2485. doi: 10.1007/s11098-016-0623-5

Petitmengin, C., Remilleux, A., and Valenzuela-Moguillansky, C. (2019). 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

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. (2010). "A neurophenomenological study of epileptic seizure anticipation," in *Handbook of Phenomenology and Cognitive Science*, eds D. Schmicking and S. Gallagher (Dordrecht: Springer), 471–499. doi: 10.1007/978-90-481-2646-0_25

Piccinini, G. (2010). How to improve on heterophenomenology: the self-measurement methodology of first-person data. *J. Conscious. Stud.* 17, 84–106.

Printz, J. (2018). "Is consciousness a trick or treat," in *The Philosophy of Daniel Dennett*, ed. B. Huebner (Oxford: Oxford University Press), 171–195.

Rosenthal, D. (2018). "Seeming to seem," in *The Philosophy of Daniel Dennett*, ed. B. Huebner (Oxford: Oxford University Press), 133–164.

Roy, J.-M. (2007). Heterophenomenology and phenomenological skepticism. *Phenomenol. Cogn. Sci.* 6, 1–20. doi: 10.1007/s11097-006-9030-2

Schmicking, D. (2010). "A toolbox of phenomenological methods," in *Handbook of Phenomenology and Cognitive Science*, eds D. Schmicking and S. Gallagher (Dordrecht: Springer), 35–55. doi: 10.1007/978-90-481-2646-0_3

Schwitzgebel, E. (2007). No unchallengeable epistemic authority, of any sort, regarding our own conscious experience – contra Dennett? *Phenomenol. Cogn. Sci.* 6, 107–113. doi: 10.1007/s11097-006-9034-y

Schwitzgebel, E. (2011). *Perplexities of Consciousness*. Cambridge, MA: MIT Press.

Sokolowski, R. (2008). Husserl's discovery of philosophical discourse. *Husserl Stud.* 24, 167–175. doi: 10.1007/s10743-008-9043-5

Thomasson, A. L. (2003). Introspection and phenomenological method. *Phenomenol. Cogn. Sci.* 2, 239–254. doi: 10.1023/b:phen.0000004927.79475.46

Thompson, D. L. (2000). "Phenomenology and heterophenomenology: Husserl and Dennett on reality and science," in *Dennett's Philosophy: A Comprehensive Assessment*, eds D. Ross, A. Brook, and D. L. Thompson (Cambridge, MA: MIT Press), 201–218.

Thompson, D. L. (2009). *Daniel Dennett*. London & New York, NY: Continuum.

Thompson, E., Noë, A., and Pessoa, L. (1999). "Perceptual completion: a case study in phenomenology and cognitive science," in *Naturalizing Phenomenology*, eds J. Petitot, et al. (Stanford, CA: Stanford University Press), 161–195.

Varela, F. (1996). Neurophenomenology: a methodological remedy for the hard problem. *J. Conscious. Stud.* 3, 330–349.

Varela, F. J., Thompson, E., and Rosch, E. (1991/2016). *The Embodied Mind: Cognitive Science and Human Experience.* Revised Edn. Cambridge, MA: MIT Press.

Velmans, M. (1991). Consciousness from a first-person perspective. *Behav. Brain Sci.* 14, 702–726. doi: 10.1017/S0140525X00072150

Zahavi, D. (2002). Transcendental subjectivity and metaphysics: a discussion of David Carr's Paradox of Subjectivity. *Hum. Stud.* 25, 103–116. doi: 10.1023/a:1014819330256

Zahavi, D. (2007). Killing the straw man: Dennett and phenomenology. *Phenomenol. Cogn. Sci.* 6, 21–43. doi: 10.1007/s11097-006-9038-7

Zahavi, D. (2015). "Phenomenology of reflection," in *Commentary on Husserl's Ideas I*, ed. A. Staiti (Berlin: de Gruyter), 177–193.

Zahavi, D. (2017). *Husserl's Legacy*. Oxford: Oxford University Press.

Zahavi, D. (2019). Applied phenomenology: why it is safe to ignore the epoché. *Continent. Philos. Rev.* doi: 10.1007/s11007-019-09463-y. [Epub ahead of print].

Zaner, R. M. (1973). "The art of free phantasy in rigorous phenomenological science," in *Phenomenology: Continuation and Criticism*, eds D. Cairns, et al. (Den Haag: Martinus Nijhoff), 192–219. doi: 10.1007/978-94-010-2377-1_12

Zawidzki, T. (2007). *Dennett*. Oxford: Oneworld.

Zawidzki, T. (2018). "The many roles of the intentional stance," in *The Philosophy of Daniel Dennett*, ed. B. Huebner (Oxford: Oxford University Press), 36–56.

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

HYPOTHESIS AND THEORY
published: 25 September 2020
doi: 10.3389/fpsyg.2020.558709

Inhibited Intentionality: On Possible Self-Understanding in Cases of Weak Agency

*Line Ryberg Ingerslev**

Department of Philosophy, Faculty of Humanities, Julius-Maximilians-Universität Würzburg, Würzburg, Germany

The paper addresses the question of how to approach consciousness in unreflective actions. Unreflective actions differ from reflective, conscious actions in that the intentional description under which the agent knows what she is doing is not available or present to the agent at the moment of acting. Yet, unreflective actions belong to the field in which an agent experiences herself as capable of acting. Some unreflective actions, however, narrow this field and can be characterized by intentionality being inhibited. By studying inhibited intentionality in unreflective actions, the aim of the paper is to show how weaker forms of action urge us to expand our overall understanding of action. If we expand the field of actions such that it encompasses also some of the involuntary aspects of action, we are able to understand how unreflective actions can remain actions and do not fall under the scope of automatic behavior. With the notion of weak agency, the paper thus addresses one aspect of unreflective action, namely, "inhibited intentionality" in which an agent feels a diminished sense of authorship in relation to her possibility for self-understanding. The notion of weak agency clarifies how agency itself remains intact but can involve a process of appropriation of one's actions as one's own. With a diachronic account of consciousness in unreflective action, the paper accounts for possible self-understanding in cases where none seems available at the moment of action.

Keywords: unreflective actions, habits, consciousness, action, responsibility, diachronicity

OPEN ACCESS

Edited by:
Wenjing Cai,
Shanghai Jiao Tong University, China

Reviewed by:
James Jardine,
University of Jyväskylä, Finland
Maren Wehrle,
Erasmus University Rotterdam, Netherlands

***Correspondence:**
Line Ryberg Ingerslev
line.ingerslev@uni-wuerzburg.de

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

Received: *03 May 2020*
Accepted: *28 August 2020*
Published: *25 September 2020*

Citation:
Ingerslev LR (2020) Inhibited Intentionality: On Possible Self-Understanding in Cases of Weak Agency. Front. Psychol. 11:558709. doi: 10.3389/fpsyg.2020.558709

INTRODUCTION

> *At any moment in any man's waking and conscious life there is always a set of possible true answers to the questions—"What is he doing now?" For human beings, to be conscious is to have active intentions. (Hampshire, 1970, p. 169).*

Which behavior deserves the status of an action or what characterizes human action is, and has been, widely debated. Many of our actions are carried out in an unthinking, unreflective way. The way we get out of bed in the morning, the way we drive to work, how we greet our colleagues; our routines and daily doings often go by without us noticing what we are doing. How do we describe consciousness in unreflective actions, and can such forms of behavior be described as actions at all? According to standard accounts in the philosophical theory of action (Bratman, 1987; Anscombe, 2001; Davidson, 2001), what constitutes an action is that it is done for a reason and that the agent knows the description under which his action is intentional. Thus, if I, unreflectively and

inattentively, put my left shoe on my right foot, this mistake does not fall under the scope of an action. The action in this case is that I put my shoes on in order to get dressed, and I am aware of dressing more or less attentively. The left shoe on my right foot is an accident; it falls outside the scope of what I am conscious of when acting.[1]

Thus, if we follow the standard theory when we consider consciousness in action, we typically want to hold on to the following assumptions, as they are both intuitively plausible:

(1) My behavior deserves the status of action, when I am conscious of doing it for a reason under some description.
(2) The behavior for which I ought to feel responsible is the behavior that deserves the status of actions.

However, if we hold on to both assumptions, some cases of unreflective action pose a problem for the standard theory of action. The problem I want to draw attention to is that the scope of behavior that I am conscious of intentionally doing under some description is smaller than the scope of the behavior for which I intuitively feel responsible. This discrepancy between what I do and what I can feel responsible for is central to many of our daily routines; therefore, it deserves theoretical attention.

The following example illustrates the discrepancy in question: Every morning I greet my colleagues upon entering our shared office; one I greet formally; another I greet in a playful tone of voice. If I was asked why that is, or if my serious colleague asked me whether there is a reason, why I treat him less playfully, I would have no good answer. Yet, intuitively, I do feel responsible for treating him as less playful or the other one as less serious, for that matter, even though I had no intention to do so. Despite the fact that I greet my colleagues for a reason, and that my action is intentional under this description, I still intuitively feel responsible for aspects of my behavior of which I am not consciously aware. I am not aware of greeting them differently, and thus, part of my behavior is something I intuitively feel responsible for despite the fact that I am acting unreflectively but for a reason under some other description.

By contrast, let us say a third colleague is in the room and I do not greet this person despite having seen her. In this case, I am responsible for acting rudely by not greeting her. Or, if I close my eyes as a reflex because of the sharp sunlight coming in from the windows, I might put my hand in front of my eyes and wave it a little. However, this is not a greeting gesture, it is a reflex. Mere behavior of this kind is without communicative intent.

According to assumption 2, action and responsibility are coextensive. If I am responsible for something, it is because it is an action of mine. According to assumption 1, action and consciousness are coextensive. That is, if something is an action of mine, I am conscious of what I am doing under its intentional description. However, according to my example, there are certain ways of doing things of which I am not consciously aware but for which I do remain responsible. To phrase it differently: In the light of habits, routines, and other aspects of my doings that are not in the foreground of my conscious awareness, standard philosophical accounts of action face the challenge that there appear to be actions of mine that I am not consciously aware of as intentional under a description but for which I remain responsible.

When responding to this challenge from within the framework established by action theory, three logically possible strategies present themselves.

(1) We can keep assumptions 1 and 2 but deny the case: We can deny that the way I greet people differently can be part of the action for which I am responsible. I greet many people differently, but what I am responsible for is that I greet them, not how I do so, unless of course this is part of my reason for doing it. I greet a former partner differently than I greet my best friend and purposefully so. However, I find it compelling that there are indeed cases as the one described above and many others, where it is important to retain responsibility despite the lack of an intentional description under which I am conscious of the action as reasonable.
(2) Another logically possible strategy consists in keeping responsibility and action coextensive, even when I am not conscious of my action. In this case, we would keep assumption 2. Hence, there will be actions of which I am not conscious in the sense described above but for which I do remain responsible. This response argues that no strong form of reflective consciousness is required for the constitution of an action. Greeting my colleagues differently is thus an action of mine for which I am responsible, despite the lack of a consciousness of a reason for doing it. Thus, we deny assumption 1 but hold that even in cases of unreflective doings, I remain responsible for my behavior. In this way, the scope of what I am consciously aware of is smaller than the behavior for which I am responsible.
(3) A third logically possible strategy consists in wanting to keep action and consciousness coextensive. My actions are coextensive with my behavior where I am conscious of the description that makes it intentional. The unreflective manner in which I greet someone is thus not part of my action. We keep assumption 1. This means that the field of actions is smaller than the behavior for which I am responsible: Unreflective doings fall outside the scope of action, but we might still be responsible for them. We thus deny assumption 2 and acknowledge that I can be responsible even for forms of unreflective behavior and not only for conscious actions.

   Both the second and third strategy take the greeting example seriously, but they do so at the expense of one of the intuitive assumptions we began with. I wish to propose a fourth strategy.
(4) I will argue that we can keep both assumptions but that they must be interpreted diachronically. In greeting my colleagues differently in an unthinking manner, I am

[1]There are of course many ways in which we are conscious of ourselves when acting. For example, I am sensorily conscious of how my body feels, perceptually conscious of my surroundings, and I am also typically reflectively conscious of what I am doing and why I am doing it. My aim in this paper is to investigate the role of the latter form of reflective consciousness as a constitutive requirement for action.

unaware of doing so, but I do remain responsible for what I do. At the time of my action, I am not aware of the description that makes my action intentional, but over time, I can come to appropriate my behavior as an action of mine by subsequently becoming aware of a description under which it was indeed intentional. Thereby, I retrospectively take responsibility for it. As long as the description under which my earlier doing was intentional can become accessible to me at a later point, this doing will be an action for which I am responsible. I thus distinguish between intention in action and the appropriation of an intention of an earlier action. The difference will be explained in the following sections.

SYNCHRONICITY

In what follows, I will look more closely at the first three logically possible strategies while arguing that a synchronicity assumption is a stake in them, that is, that the necessity of synchronicity of action consciousness and the action itself is assumed. In the case, where we keep both assumptions 1 and 2, one must deny that my manner of greeting—unless it is part of the description under which my action is intentional—can be part of my action. The agent must know the description under which an action is intentional: I am going to work (independently of whether I am attentive to the route); I am thinking about a job interview (independently of whether I bite my nails). Importantly, an intention cannot be ascribed *post hoc*; it has to be involved as a reason for one's action. Thus, nail biting and the route I follow are mere happenings. Habitual actions that are characterized by such involuntary aspects fall outside the scope of what characterizes us as agents, of what constitutes our practical identities, and ultimately of what we are responsible for as agents. The habit of talking to myself while typing is an activity that is not an intentional action of mine (Setiya, 2017). Without the possibility of first-personal avowal or the acknowledgment of authority of my doings (Moran, 2001), they fail to be actions of the kind that constitutes my practical identity. In such cases, I will typically have merely attributional knowledge of my state of mind, based on observation, and mediated by some identifying description.[2] In some cases, I might be immediately aware of my behavior yet lack the possibility of avowing what I am doing. However, in such cases, I would feel self-alienated rather than aware of myself as an agent (ibid., p. 33).

The problem with this strategy is that we are forced to narrow the field of action too far beyond our intuitive understanding of what counts as behavior for which I am responsible. To avoid this, some have suggested that we loosen the connection between reflective consciousness and action, as described in option (2) above. They have argued that we should still ascribe the status of actions to some of the behavior that we are not conscious of doing for a reason (Hursthouse, 1991; Steward, 2009; Owens, 2017). On this view, many of our actions we do unreflectively, but still, they remain actions.

A similar strategy has been pursued in the phenomenological tradition, where the link between consciousness and action is maintained by weakening the notion of consciousness required for behavior to count as action. The notion of embodied, operative intentionality serves to bring behavior within the scope of consciousness without requiring the stronger form of consciousness of intention under a description that is typically pursued in the analytical tradition. In my practical engagement in dancing, hammering, and walking, I embody intentionality, rather than reflectively carrying out a conscious action plan. Or, to put it differently, consciousness is broadened such that it not only refers to reflective consciousness in the sense that I know the description under which my action is intentional or done for a reason. Rather, it also includes forms of unreflective, embedded, and enactive consciousness. According to Merleau-Ponty, for instance,

> *Consciousness is originarily not an "I think that," but rather an "I can" [. . .] Consciousness is being toward the thing through the intermediary of the body. A movement is learned when the body has understood it, that is, when it has incorporated it into a subject's "world," and to move one's body is to aim at the things through it, or to allow one's body to respond to their solicitation, which is exerted upon the body without any representation. (Merleau-Ponty, 2012, pp.139–140).*

As Merleau-Ponty writes: "My body has its world, or understands its world without having to go through 'representations,' or without being subordinated to a 'symbolic' or 'objectifying function'" (ibid., p. 141). This means that when the body acquires a habit, it comes to understand something in the world, and this understanding is practical: "This formula will seem absurd if 'understanding' is the act of subsuming a sensory given under an idea, and if the body is a mere object." (ibid., p.146). Merleau-Ponty's notion of habit clearly differs from any notion that would reduce habit to an automatic reflex, a tic or an otherwise involuntary side effect. Rather, habits are a form of practical, embodied understanding: "To understand is to experience the accord between what we aim at and what is given, between the intention and the realization—and the body is our anchorage in a world" (ibid., p.146). The affordance character of objects, situations, and other people (ibid., pp. 191–192) is perceived through the lived body, and the responsive answering of these calls is experienced as an embodied capacity. The "I can" is an experiential structure that shapes our bodily existence. Understanding is an embodied practice that precedes theoretical knowledge. In this way, embodied intentionality is operative beneath our consciously minded actions. Employing a broader conception of consciousness in action in this particular manner results in non-reflective embodied engagement figuring as a form of action. In such a theory, the domain of action is bigger than the domain of reflective, conscious doings. The focus is on the agent's engagement and on embodied intentionality; hence, in terms of a theory of action, this means expanding the field of action such that even unreflective forms of behavior count as actions.[3] Within analytical philosophy, we find similar arguments

[2]In Section "*Objection: Does Diachronicity Exclude First-Personal Authority?*," I return to Moran's conceptions of first-personal authority and avowal.

[3]In what follows, I shall propose a fourth argumentative strategy that draws on aspects of the Merleau-Pontian phenomenology (see sections "*Inhibited*

that reflective consciousness is not necessary for behavior to count as action. As Steward (2009) argues, it is the first personal capacity to do otherwise, which settles whether something is an action or not. When doing something unreflectively such as fiddling with one's jewelry, it is one's capacity to do otherwise that settles whether it is an action. She thinks that we should not overmentalize what constitutes an action. Rather, the two-way force consists in the ability of doing or refraining from doing, and it constitutes whether some behavior is an action (ibid.). If an agent could not have done otherwise, what she did was not an action. The question of what it means to have the ability to do otherwise, to do something or refrain from doing it, is even intact in cases that put free will under pressure, as Pickard (2015) argues in the context of substance abuse and addiction. The addict is someone whose ability to do otherwise remains intact. The addict's actions of consuming and using substances is something she can refrain from, and thus, the addict is *not* compelled to use drugs, although the addict's capacity to do otherwise is weakened. She is *not* an unfree agent. For Pickard:

> *Our common sense conception of agency draws a basic distinction between actions and mere bodily movements, such as automatic reflexes. What makes a piece of behavior an action, as opposed to a mere bodily movement, is that it is voluntary, where this means that the agent can exercise choice and at least a degree of control over the behavior. [. . .] [O]ne is responsible for actions, as opposed to automatic reflexes, because it is up to one whether and how one acts. (Pickard, 2011, 212).*

The consequence of Pickard's view is that we can hold the addict and other agents suffering from "disorders of agency" (Pickard, 2015) responsible for their doings because they are not compelled; they can and could have done otherwise. This means that even in cases where agency seems diminished and weakened, we still hold an agent responsible because she can do and could have done otherwise. Instead of blaming her, we hold her responsible (ibid., pp. 140–2).

For theories such as the ones just mentioned, the greeting example is an action of mine for which I am responsible, even though I am not conscious of it in a strong sense that requires knowledge of the intention with which it was done. Rather, I am responsible because we have expanded the field of action such that it is broader than the scope of reflective consciousness. This expansion allows us to keep action coextensive with responsibility.

As for the logical possibility where action and consciousness are kept coextensive, but it is denied that action and responsibility are so, we would have to imagine a theory that denies our intuitions about which behavior we should feel responsible for. Here, the attempt is to argue that there is no problem with thinking we are responsible in the greeting example despite the fact that the way of greeting is not part of my action. This solution would thus preserve the tie between action and consciousness but at the cost of sacrificing our intuitions about which behavior we should feel responsible for.

What we see is that we can either (1) keep both assumptions synchronically, but then the greeting case must be denied. We can (2) deny that reflective consciousness is coextensive with action and accept that the scope of action must be broadened in order to keep the link between action and responsibility, or (3) we can deny that action and responsibility are coextensive and argue that we can be responsible for behavior that is not action. Thus, either we cannot explain the greeting case, or we are forced to give up at least one of the intuitive assumptions often thought central to action theory.

I wish to propose a fourth option that preserves both assumptions mentioned above. My aim is to do justice to our intuitions about which behavior we should feel responsible for. The crucial step of my argument is to deny an underlying assumption that has been governing all sides of the debate so far. This is the assumption that when we evaluate the status of some behavior to determine whether it is an action, then it is solely the contemporaneous consciousness of the individual we need to examine. Instead, I want to argue that only as far as I *can* become conscious of some description under which the action was intentional does my behavior deserve to be called an action. This account entails that some actions of mine can return to me as questions, to which I can only appropriate the reply diachronically, over time.

In the discussed logical strategies (1)–(3), we find what I will call the synchronicity assumption to be operative: The states of consciousness relevant to determining whether some behavior is an action are only those states that are synchronic or contemporary with the behavior. Synchronicity refers to the theoretical role of immediate first-personal insight into the intention of one's action (see also Ingerslev, 2020). There is a tendency to identify as a hallmark for agency the simultaneous relation between one's first personal insight into one's intention and the action being performed. That is, if consciousness of intention is relevant in order to classify whether something is an action or not, then it is synchronically relevant. However, we do not always have immediate insight into the intention of our habitual doings and, in some cases, can only gain this insight over time.

Either it is required that consciousness of the intention is decisive for whether something is an action. Therefore, habitual doings where one lacks insight into one's intention are not actions, or habitual doings are considered actions, but then knowledge of one's intention is not required for something to be an action. Practically, this means that upon asking an agent what she is doing, either she would immediately know the intention with which she is doing something or she would understand more broadly which practical activity she is engaged in, although she is performing it unthinkingly. Thereby, the possible awareness of habitual actions is synchronically related to the possibility for self-understanding. In this way, full self-understanding afforded

Intentionality As an Expression of Weak Agency" and "*Objection: Does Diachronicity Exclude First-Personal Authority?*"). I shall, however, emphasize the first-personal agential insight in temporally expanded agency and thus endorse Merleau-Ponty's analysis of confined freedom while accepting the framework offered by action theory. If we endorse merely the expansion of the domain of action to also encompass unreflective doings, we cannot spell out the possible self-understanding of an agent of her own doings over time. I thus attempt to spell out the action theory needed for engaging with the temporal aspects of agency that Merleau-Ponty does not himself address in his use of a psychoanalytic example, see Section "*Objection: Does Diachronicity Exclude First-Personal Authority?*"

by an agent's awareness of her habitual behavior is immediately available by the time of the awareness. The general tendency expressed by the synchronicity assumption is at play in all the positions presented provided above.

In the case of unreflective manners of greeting, we can thus either argue (1) that it is not part of my action because, synchronically, I do not know of an intention that would make this behavior intentional; (2) unreflectively greeting in this particular way is an action of mine because, synchronically, I am embedded and enactively involved as agent also in unreflective cases. I could, synchronically, have refrained from greeting this or that way; I was not compelled. (3) Synchronically, I am unaware of my doing, but I remain responsible as the action is ascribed to me at the time T.

By contrast, I wish to argue that it is by a process of appropriation of our intentions that we are able to understand some of our unreflective doings as actions. The description under which my action is intentional is still relevant to determine whether some behavior of mine is an action. However, it is not the synchronically accessible description but rather those aspects that I *can at some time T* become conscious of as belonging to a description under which my action was done intentionally. Diachronically, therefore, I can assume ownership of my doing in such a way that I understand my actions retrospectively as belonging to me in a stronger sense. I take on the responsibility of having done them for a reason. I have been responsible all along, but I come to assume ownership and responsibility for my action over time. Rather than my greeting behavior being an accidental aspect of my action, I can appropriate it as something I did; I have come to realize that I do greet my colleagues differently for a reason. Maybe, without having thought about it explicitly, I do think that my one colleague prefers a serious work ethos and therefore I greet him in a more formal way. What characterizes behavior of mine that I can come to appropriate over time as an action of mine is that this behavior can come to matter rationally for my self-understanding. Instead of providing retrospectively a causal explanation of something I did, I provide a rational account that matters for how I see myself as a person. If asked why I keep being distant and snappy at my good friend, I might come to realize that I have been angry with her for a long time. If I believe that due to stress and a heavy workload I could not have acted differently, I would have explained my aggressive tone causally. By contrast, appropriating my actions diachronically means to engage with my behavior and try to realize for how long I have acted like this and, in responding to these doings, to understand them as something done by me for a reason. This might lead to doubts whether responsibility can be thus construed retrospectively and to the objection that forms of antirealism concerning action will be unavoidable. I will return to these issues in Section "*Objection: Does Diachronicity exclude first-personal authority?*"

Repetitive behavior of a certain opaque and incomprehensible kind is at the same time something I can come to understand as actions of mine. Whether I keep smoking, keep greeting my friend seriously, or whether I might have been in love with someone for a long time without realizing it, these cases entail behavior that I can diachronically come to realize that I did for a reason. Therefore, I can appropriate them as actions of mine and take responsibility for them. Whereas the cases and consequences will differ between smoking out of habit, greeting my colleagues differently, bullying or discriminating against someone, the structure of how we diachronically assume responsibility for our past behavior is the same. What I aim to show in this paper is that there is such a thing as appropriating one's past behavior as action.

By realizing that the scope of conscious action for which I am responsible also encompasses cases where the consciousness of action occurs diachronically, we can account for how it is that self-understanding is important *even* in cases of unreflective actions. If we accept synchronicity, many of our daily routines and other involuntary aspects of our unreflective doing are in danger of being out of our conscious reach. That is, a consequence of denying that we can diachronically appropriate behavior as actions of ours is that we are unable to account for how we increase our understanding of our past agency by reflecting on our reasons for earlier behavior. Without this diachronic option of appropriation, my self-understanding becomes opaque, blocked, or even barred. Ultimately, self-estrangement can be a result of not accepting and acknowledging diachronicity as a process of appropriating one's own habitual doings as actions.

The field that opens with diachronicity is larger than what I aim to address in this paper. My aim is to focus on cases where part of my emotional response to certain situations is beyond the synchronically available description that makes my action intentional. However, with a diachronic perspective, some of my habitual doings can be appropriated, and what was blocking or barring an adequate account of my self-understanding is opened up. In what follows, I will look into a case of inhibited intentionality and propose that we understand it as a case of weak agency. This will help us address the challenge for action theory, namely, the status of unreflective action as personal.

INHIBITED INTENTIONALITY

In this section, I will argue that those unreflective actions that matter for our self-understanding and can be appropriated diachronically are not merely of a peripheral kind, such as the greeting case. Many of our daily patterns of behavior, for example, those that are socially and culturally shaped, fall within the scope of unreflective action.

Many involuntary aspects are automatic and can play no role for my self-understanding; that is, they do not fall under the scope of what is personal. Consider how many times I blink per minute or the size of personal space measured in relation to how I place objects belonging to me when traveling by train. These are aspects of my actions that I do involuntarily and with little sense of agency involved in them. However, some involuntary aspects of my actions do play an important role for my potential self-understanding as an agent. Maybe I have to look down when talking to certain persons, maybe I cannot use my body normally when I throw a ball, or maybe I cannot sit far away from my personal belongings and have to cling to my purse in order to feel safe in a public train. These involuntary aspects are expressed

in my bodily habits. In order for me to understand them as belonging to me in a stronger sense, that is, as something I do and that I am responsible for doing, the diminished sense of agency at stake in these forms of action must be addressed. Such doings of mine play an important role for my self-understanding. If I keep avoiding certain persons, if I feel unsafe without my personal belongings clinging to me, or if I greet my friends differently, I must know why I do so in order to understand my own actions as mine and to understand who I am as a person.

In what follows, I will take up Iris M. Young's example of throwing a ball in a certain way. In her example (Young, 1980), she addressed the inhibiting effect that cultural education can have on a person or a group of persons and their possibility for self-expression and self-understanding. In the case of throwing like a girl, she makes the philosophical point that embodied intentionality can be inhibited by cultural upbringing. Whereas this is less surprising, I will focus on her theoretical explanation of how such inhibition is experienced. I am interested in the example in relation to the question of how the involuntary aspects of our habitual doings can still be appropriated as something done by us, which makes inhibited intentionality a case of weak agency.[4]

In Iris Young's 1980 paper, she proposes that embodied intentionality can be experienced and enacted as inhibited. Young discusses the role of gender by questioning how we embody intentionality. She claims that—in a specific culture, at a specific time in history—throwing, standing, walking, talking, sitting, and laughing is experienced differently relative to one's gender. Young does not make an exhaustive claim, but her point is that the ingraining and bodily habituation of gender stereotypes through various social processes come with forms of inhibited intentionality for women. In order to assess this claim, we need to make certain preliminary considerations. For Husserl, the notion of bodily awareness is characterized by an experiential "I can." This "I can" neither is a belief nor is it experienced consciously as a propositional truth. Rather, embodiment is enacted and prereflectively experienced under the condition of an "I can," understood as practical possibility (Husserl, 1989, p. 159 ff., p. 165 ff., p. 269 ff.). Another way to phrase this is that the way we experience embodiment prereflectively is as the capacity to do. . .. Or the ability to. . .. Being embodied means to be intentionally directed toward the world and to feel moved by the world, i.e., by objects, people, and situations. Movement, action, and activity are thus structurally characterized by an embodied ability to. . .. Further, this is a way in which we embody consciousness: The prereflective awareness of "I can" is constitutive of how we embody intentionality; we are directed toward the world as embodied beings. As we said above with Merleau-Ponty, the affordance character of objects, situations, and other people is perceived through the lived body, and the responsive answering of these calls is experienced as an embodied capacity, the "I can" as an experiential structure that shapes our bodily existence. Practical understanding is an embodied practice that precedes theoretical knowledge and propositional attitudes. We inhabit a world practically before we understand it theoretically. In this way, embodied intentionality is operative beneath our consciously minded actions. To put it differently, the experiential structure of embodiment is an "I can" that mediates our bodily movements and our comportment and is prior to representational, theoretical understanding.

Young investigates the idea that women comport themselves differently from men and illustrates this by the example of throwing a ball:

> *Women tend not to move out and meet the motion of the ball, but rather tend to stay in one place and react to the ball's motion only when it has arrived within the space where she is. The timidity, immobility and uncertainty which frequently characterize feminine movement project a limited space for the feminine "I can." (Young, 1980, p. 150).*

Young argues that there is something like feminine bodily existence where the embodied intentionality is experienced as inhibited in the sense of an "I cannot." The idea is that the "I can" remains fundamental for our embodied existence but that it can be modified as we take on certain cultural life forms. Space is culturally shaped and coded such that certain groups are allowed to move in certain ways as they follow the forms and norms of collective education.

What Young refers to as feminine bodily existence neither is meant to be exhaustive nor is it meant to be universal (ibid., p.139). Her account "claims only to describe the modalities of feminine bodily existence for women situated in contemporary, advanced industrial, urban, and commercial societies" (Ibid., pp. 139–40). That is, she specifically targets a kind of comportment that is set in time and space, which she terms feminine bodily existence, and she seeks to describe its phenomenological structure.[5] I believe the strength of this account lies less in how it describes feminine bodily existence and more in how it sheds light on the possibility that agency can be weakened despite the fact that the agent is free. With Young's example, it is possible to expand the field of possible involuntary aspects of one's doings from our local greeting example to the more global case of cultural life forms under which certain groups suffer. I believe that Young did not develop the potential of her account in the broader field of philosophy of action. The notion of inhibited intentionality is fruitful for our understanding of the relation between involuntary aspects of agency and the possibilities for self-understanding.

For a subject that experiences feminine bodily existence, embodied intentionality is inhibited. According to Merleau-Ponty, the lived body structurally describes how subjects are embedded in and belonging to the world. To embody an "I can" means to be intentionally directed toward this world by being capable of. . .. Normally, the lived body is the unifying synthesis

[4]I will focus specifically on the action theoretical question that Young's influential and highly debated paper raises. This means that I will leave out important discussions of Young's own assumptions and of the implications of her work.

[5]In what follows, I will refer to "feminine existence," "she," "her," "woman," and "womanhood," etc. as does Young. That means, independently of whether or not one believes such a thing to ontologically exist or to fit the descriptions given by Young, the theoretical claim I am interested in is that a modification of the underlying "I can" that characterizes embodied intentionality is possible and that this modification is described by the term "feminine" or "woman".

of our experiences. For Merleau-Ponty, a bodily synthesis refers to how the lived body ensures and enables the unity of embodied perception. When I see a cup, I experience the synthesis of my perceptual impressions, not through my mental representation of a cup, but because of the primordial rootedness of perceptual affordances in my lived body. My perceptual act is an embodied comprehension of the cup:

> *[T]o habituate oneself to a hat, an automobile, or a cane is to take up residence in them, or inversely, to make them participate within the voluminosity of one's own body. Habit expresses the power we have of dilating our being in the world, or of altering our existence through incorporating new instruments. (Merleau-Ponty, 2012, pp. 144–5).*

This means that "[c]onsciouness is being toward the thing through the intermediary of the body" (ibid., p. 140). Here, consciousness is embodied in such a way that every perceptual act is rooted in bodily practical understanding. Thus, what is embodied in practical understanding is at the same time an incorporation of one's world. As Merleau-Ponty phrases it: "A movement is learned when the body has understood it, that is, when it has incorporated it into its 'world,' and to move one's body is to aim at the things through it, or to allow one's body to respond to their solicitation, which is exerted upon the body without any representation" (Ibid., p. 140). The body is not "in" time or "in" space, but inhabits times and space as an active linking them together. Thus, "[i]nsofar as I have a body and insofar as I act in the world through it, space and time are not for me a mere summation of juxtaposed points, and no more are they, for that matter, an infinity of relations synthesized by my consciousness in which my body would be impacted" (Ibid., p. 141). The synthesis of the body is thus that of having a world, or "understanding its world without having to go through 'representations,' or without being subordinated to a 'symbolic' or 'objectifying function'" (Ibid., p. 141). In the case of feminine bodily existence, the embodied synthesis is disrupted. For someone who exists in this feminine way, the unifying synthesis provided by the lived body is blocked; thus, the experiencing subject remains stuck in not expressing herself fully and not receiving her world fully. She is detached in some aspects from worldly belonging. She does not manage full worldly transcendence, and she experiences herself—through her embodied existence—as someone who feels less capable, less open, less powerful, more insecure, and more concerned with her own fragility (Young, 1980, pp. 143–4); she is held back in immanence (ibid., pp. 144–5). The world for a woman is constituted as a field of inhibition rather than as a world to be inhabited. Further, she experiences embodiment as a something that is not tied to fields of possibilities because these possibilities in fact occur as possible for someone else than her; she is caught in the lived absence of possibilities and thus in lived inhibition. Experientially, the field that opens with an embodied "I can" is closed off for her, and as a field of inhibition, it does not offer her the means of self-expression; she experiences this field as indicative of her own incapacity, i.e., her embodied intentionality is experienced as inhibited (ibid., p. 147).

> *Typically, the feminine body underuses its real capacity, both as the potentiality of its physical size and strength and as the real skills and coordination which are available to it. Feminine bodily existence is an inhibited intentionality, which simultaneously reaches toward a projected end with an "I can" and withholds its full bodily commitment to that end in a self-imposed "I cannot." (Ibid., p. 146).*

Granted that we are all limited in our bodily capacities, the experienced "I cannot" belongs just as much to the nature of embodied intentionality (Ingerslev, 2013). The case Young is making is that for feminine bodily existence, the experienced "I cannot" is self-imposed: "When the woman enters a task with inhibited intentionality, she projects the possibilities of that task—thus projects an 'I *can*'—but projects them merely as the possibilities of 'someone,' and not truly *her* possibilities—and thus projects an 'I cannot'" (Young, 1980, p. 147).

This point is crucial for our present task of arguing that inhibited intentionality is an aspect of weak agency that can be appropriated diachronically. The experience of a self-imposed "I cannot" shows an aspect of who I am and what I do that might at the moment be something that I am not conscious of. I do not know that I inhibit myself, but it remains something for which I am responsible and that I can come to realize—maybe upon rumination, critical thinking or therapy—as something I did to myself. This is where the notion of diachronic self-understanding comes into play. It is not a question of taking more sports classes in order to come to throw more fully. Rather, it is a matter of appropriating an embodied worldview, a practical self-understanding, as something I have acted under and that I might want to change. If we could not diachronically come to realize our own reasons for action, then our theory of action cannot explain why today, upon acquiring this consciousness, I should feel responsible, or importantly characterized as an agent, by my own past self-inhibition.

Young's claim is that feminine bodily existence isolates movements and does not make use of the full bodily potential to perform an activity. As illustrated by the example of throwing, a woman might only use her arm in throwing a ball. By comparison, a non-feminine bodily existence would turn the upper body, use the strength of a firm grounded position, and the other arm in aiming, etc.: "The undirectedness and wasted motion which is often an aspect of feminine engagement in a task also manifests this lack of body unity" (ibid., p.147). Not only is there an undirectedness or a waste of motion in the sense that some movements could be more focused and fully executed, but this is the superficial part of the problem. The real problem is the lack of bodily unity. In reaching for the objects, grasping, moving, and perceiving, the female body is only practically set in motion; a woman does not carry her movements fully through. Thus, no bodily synthesis can be fully made. This means, in its widest consequence, that the bodily synthesis is disrupted and that forms of self-estrangement and derealization are part of feminine self-understanding. To see why that is, we must return to the synthesis of the lived body, as Merleau-Ponty understands it. As said above, the embodied synthesis anchors my experiences, and it provides the background for my self-understanding altogether:

> *[T]he consciousness that I have of it [my body] is not a thought, that is, I cannot decompose and recompose this consciousness in order to form a clear idea. Its unity is implicit and confused. It is*

> *always something other than what it is: always sexuality and at the same time as freedom, always rooted in nature at the very moment when it is transformed by culture, it is never self-enclosed but never transcended. [. . .] Thus, I am my body, at least to the extent that I have an acquisition, and reciprocally my body is something like a natural subject, or a provisional sketch of my total being. The experience of one's own body, then, is opposed to the reflective movement that disentangles the object from the subject and the subject from the object, and that only gives us thought about the body or the body as an idea, and not the experience of the body or the body in reality. (Merleau-Ponty, 2012, p. 205).*

When the bodily synthesis is disrupted and self-understanding is realized as derealization and self-estrangement, the experiential field itself is experienced as inhibiting, threatening, and possibly even foreign. What could have been part of a world-for-me is turned into a foreign field that alienates me from myself. As a result, feminine bodily existence is discontinuously realized: "feminine bodily existence stands in discontinuous unity with itself and its surroundings" (Young, 1980, p. 147). Overall, feminine bodily comportment is characterized as the failure to make full use of the body's spatial and lateral potentialities (ibid., p. 142). When I experience myself as bodily inhibited, no bodily synthesis can be fully reached. As a result, my self-understanding is equally shattered and disrupted since I cannot understand why I feel and behave this way. The relation between my movements and my world, my movement and reality, is in this way disrupted and inhibited.

INHIBITED INTENTIONALITY AS AN EXPRESSION OF WEAK AGENCY

In this section, I will introduce the notion of weak agency with the aim of specifying what is characteristic for unreflective and habitual behavior that can be appropriated as actions of mine. If we take up the idea of feminine bodily existence as described by Young, it is a technical term for how I act under internalized superimposed structures. The way I throw, walk, or address people differently is part of my unreflective actions. I might not be consciously aware of how I walk, and while strangely unconscious of my gait, it remains my way of walking. This aspect of my bodily habits calls my agency into question (Ingerslev, 2020). If I am not aware, say, of how I cringe in front of certain authorities (Freud, 1914; Lear, 1998), how can I assume ownership of these doings of mine as actions? If I am acting out fearful and inhibited behavior or emotional traumas (Freud, 1909), how do I come to appropriate these actions as mine; how do I take responsibility for them? What is special about the involuntary aspects of inhibited intentionality is that they are tied to our self-understanding while being temporally beyond our control. This means that something I do that involves inhibiting my field of action is at the same time crucial for my possibility for self-understanding. If I am the one throwing the ball in an inhibited way, or cringing in front of authoritative persons, I must know how these doings belong to me in a stronger sense. The self-imposed inhibition entails an inkling question: how do I appropriate my doings as actual actions of mine?

We might live a whole life without knowing about the depth of cultural influence on our behavioral pattern. We might remain ignorant of the many layers of body memory that affect our ways of responding to people and situations (Fuchs, 2012, 2018). However, what the notion of weak agency allows us to account for is the coincidence of personal habits with lived forms of self-estrangement that are even at times self-imposed. If we take Young's insights further, the idea is that these involuntary aspects of habits can shape a life form that disrupts one's self-understanding and makes it difficult to endorse one's actions as one's own; the latter remain foreign in nature to the agent, and as a result, the agent's self-understanding becomes distorted.

The technical term *Weak Agency* aims at specifying forms of unreflective and habitual doings that can be appropriated over time as one's own. By reference to Young's claim that intentionality can be inhibited and, further, can be taken on as a self-estranged life form, the notion of weak agency pursues the possibilities of self-understanding within forms of behavior that otherwise seem less agential or less free. Limit cases of weak agency where hardly any agential freedom is at stake can be found in life forms where the self-estrangement and the objectification is close to total. Such cases entail less of an opening for transformation and appropriation over time (Honneth, 2008). When agency is no longer simply weak but blocked, a certain life form is destructive and cannot be appropriated as in the commodification of bodies or dehumanizing reification of human lives mentioned by Honneth. We can think of cases of hierarchical or religious indoctrination. In such cases, the only diachronic understanding of my past behavior available is that someone else made me think or act in a certain way. I cannot come to appropriate that I had a reason for behaving the way I did. For whatever reason there was for my behavior, it was not a reason of mine but of those who indoctrinated me. Habitual behavior differs from these limit cases in that we find within the habits themselves an opening for appropriation of one's own behavior. The openness of one's own habits to transformation is tied to our possibility for self-understanding. Diachronically, I can come to take responsibility for the inhibited intentionality that shapes my unreflective doings, and thereby, I can come to appropriate them as actions of mine. Iris Young's case of feminine existence is thus an example of a life form that can lead to self-estrangement, but it also entails an opening for appropriation; it is a case of weak agency.

The reason why involuntary structures are interesting for our understanding of weak agency is due to the tension between a diminished sense of control and the intimate familiarity in bodily habits. We want to understand how the involuntary aspects of habit are more than just impersonal happenings, as they belong to me in a strong sense; I am the person who is acting freely, yet I am involuntarily inhibited in my bodily existence. I am the one acting under superimposed internalized structures, yet I am also the one who can take responsibility for this doing diachronically and thereby come to understand something about myself. The notion of weak agency thus captures that the degree to which the agent experiences herself as the author of her own actions can differ, i.e., the degree with which an agent feels in

charge of her own actions can vary. When this degree is low, the agent might not synchronically sense herself as the author of her actions in a strong sense. She might know explicitly that she is doing something or that she often does this or that, but she might not intend to do so or want to do so. The weakness at stake can thus be characterized as unwilled, as inattentive or unfocused doings, or even as an involuntary act, not because the action is forced but because it occurs independently of the agent's contemporaneous self-conscious understanding of her actions. By contrast, strong agency is tied to cases where we deliberatively and with reason perform certain kinds of actions; we consciously authorize a certain doing, and we are capable of providing the reasons of our action. I decided to take the job, or I realized our friendship could not continue; in these cases, a process of deliberation leads to a decision and culminates in action. In taking the job or ending a friendship, the agent senses agency as authorship in a strong sense; this is a doing of mine. It is not difficult to see how such doings play a central role for our self-understanding. I want to be the kind of person who is a good friend in this particular way; this is not possible with Y, and thus, I have to end the friendship. Integrity is one way to describe such a relation (Korsgaard, 1996; Crowell, 2013). A strong agent is someone who, upon consideration, provides reasons for her action: Do I want to go out for a beer and be a good colleague, or, do I want to pick up my son and take him to football training (Crowell, 2013)? The measurement for my integrity is the normative source of my actions. I take responsibility for wanting to be a good colleague, prioritizing collegial chat over a rainy day at the football field.

What is weaker in the cases I am interested in is the kind of self-understanding of what one is doing that is available to the agent presumably acting. In cases of weak agency, an immediate response to the question why are you doing X is not available to the agent. However, the response can be appropriated diachronically. Weak agency is a term that covers the remaining possibility for appropriating the full scope of one's actions as one's own. Weak agency differs from ignorance in that I can come to be aware of cringing and throwing like a girl, but I cannot immediately change it. Ignorance would be the case, where the scope of one's actions relies on epistemic barriers, not on inhibited or blocked body memory. If someone tells me that my favorite chocolate brand is run by an evil company that exploits children and women in the third world, I will stop buying it. If someone tells me, I throw like a girl, I would have to appropriate my whole being over and over again while committing to the field of my possibilities given my history while at the same time becoming the person I am. This means that the freedom involved in cases of weak agency is confined. Whether I throw like a girl or not, whether I address my embodied inhibition in therapy, these are ways in which I embody a confined freedom; I commit myself with the sedimented bodily history I have to become the person that I am—over and over again.

Merleau-Ponty reflects on this kind of confined freedom and its relation to a commitment to self-understanding by referring to the therapeutic relation in psychoanalysis. The example serves the purpose of illustrating what is meant by diachronic appropriation characteristic of weak agency:

> *By taking up a present, I* again *take hold of my past and I transform it, I alter its sense, I free myself and detach myself from it. But I only do so by committing myself elsewhere. Psychoanalytic treatment does not heal by provoking an insight into the past, but by first relating the subject to his doctor through new existential relations. "Merleau-Ponty, 2012, p. 482 (Emphasis added by the author)"].*
>
> *. . .[I]t is a question of re-living the past as signifying this or that, and the patient only achieves this by seeing his past from the perspective of his coexistence with the doctor. The complex is not dissolved by a freedom without instruments, but rather is dislocated by a new pulsation of time that has its supports and its motives. The same is true for all moments of insight: they are actual if they are sustained by a new commitment. (Ibid., p. 482).*

The notion of commitment at stake here is one that refers to possible self-understanding rather than that of a onetime promise or resolute decision. It refers to what Lear calls a dreamlike engagement, where the proper meaning of some behavior is not fixed and not identical to the manifested one (Lear, 1998 p. 97 ff.; Lear, 2017, p. 102, see also Merleau-Ponty, 2006, pp. 177–179). It is not the case that I once and for all *decide* to not throw as a girl or not cringe in front of authorities, but I commit to appropriating this behavior of mine, which will take the shape of recommitment, something I will have to do over and over again. The process of appropriation might involve several attempts to make sense of various happenings over time, cringing, not cringing, being fearful, trying not to be, etc. The commitment over time to the quest for self-understanding might at some point allow me to not cringe. Even if I keep cringing despite myself, I might work with these involuntary aspects of my actions as part of my appropriation and conscious quest for self-understanding. Commitment means that I aim for self-understanding and strive to gain insight into my reasons for acting; only the description under which my action is intentional might not always be synchronically accessible to me.

The appropriation of one's unreflective doings thus differs from being resolute, making up one's mind or having enough will power to change one's habitual behavior. Rather, it consists in the attempt of coming to terms with weaker forms of agency for which I nonetheless take responsibility. This is why Merleau-Ponty's description of confined freedom is tied to the example of commitment in psychoanalysis. In therapy, the agent addresses the involuntary aspects of her embodied personal history in order to commit herself anew as a free but weak agent. We can understand this kind of commitment as a diachronic appropriation of one's past doings *as* actions for which one take responsibility by taking them up as *part of* one's history. Thereby, the agent commits herself to striving for self-understanding. She comes to rediscover her past as a possibility for future self-understanding:

> *Freedom lies in the re*discovery *of my habitual past as* a reservoir *of possibilities, indeed, as a vigorous force actively shaping my future at every moment. It lies in our ability to enter into this force, both past and futural, intrinsically rigid and intrinsically flexible, with the stance of one who approaches the world as a place where meaning grows. "Talero, 2006, p. 203 (Emphasis added by the author)".*

What Talero describes as intrinsically rigid and at the same time intrinsically flexible is the nature of our bodily habits and unreflective actions. The involuntary aspects of these doings remain a part of the field that actively shapes our future, and as such, it can be appropriated (Ingerslev, 2020), however, diachronically, over time, as Merleau-Ponty describes. The process of approaching the involuntary aspects of one's doings and projecting them toward the world is the process of striving for self-understanding in cases of weak agency. That freedom is confined means that it does not exist outside a historical past but that this past must be appropriated over and over again.

Importantly, appropriation is not similar to being concerned with finding a set of lost intentions that we then take on, through an external third-person view on our own lives that conveys a certain useful meaning. I will return to this point in Section "*Objection: Does Diachronicity Exclude First-Personal Authority?*." It is not the case that I rediscover in my past aggressive behavior that I was indeed upset and angry with my good friend, and that is it. As Merleau-Ponty argues, it is a matter of an existential commitment that unfolds over time. While emphasizing how the involuntary aspects of our habitual doings can be diachronically appropriated, the point is not that we can simply attribute an intention to our former behavior, as if rewriting our personal history. Rather, the structures of body memory, inhibited intentionality, and emotional trauma entail an openness that returns to us as a question for our self-understanding: Why am I doing this, or why did I do it again? This opening is constitutive of weak agency, and it enables me to commit again to making sense of what I am doing as a person. The worry expressed here concerns the processual aspect of appropriation. Whereas realizing that I was upset with my good friend provides me with a reason for my past behavior, the attempt to take responsibility for my actions might entail several attempts at appropriation and at accepting and endorsing my weak agency. This is the difference spelled out between an existential commitment and the attempt of finding a reason. The former is an ongoing process that defines me as a person that actively engages with my personal history. The latter could be a case of ascriptivism, where I am unconstrained by past affairs in which intentions I ascribe to myself to explain my action. The example of a therapeutic relation is helpful in order to illustrate how we will not come out as strong agents and how we will not get rid of all of our habits. Rather, we might learn why we repeat certain patterns of behavior. The therapeutic questioning of one's behavior might help us to gain a richer self-understanding. It might help us to appropriate our past behavior as something we did intentionally and something for which we are responsible. Weak agency involves repetitive patterns of behavior that we will not as such get rid of simply by finding the reason why we repeat them, but what we do gain is insight into our own weak agency. Weak agency thus entails an important possibility for self-understanding; however, in order to see this, we needed to spell out the diachronic relation between action, consciousness, and responsibility in unreflective actions and in cases of inhibited intentionality.

OBJECTION: DOES DIACHRONICITY EXCLUDE FIRST-PERSONAL AUTHORITY?

Obviously, much remains to be explained if we accept the notions of weak agency and diachronic appropriation. What are the temporal limits to what can be appropriated; what are the epistemological constraints on what can be appropriated; what is the interrelation between memory and self-understanding over time? Does diachronicity imply antirealism about actions, that is, can any past behavior of mine be appropriated over time as an action of mine? In what remains of the paper, I will briefly discuss whether the case of diachronic accounting implies seeing oneself from a third person-perspective and whether it implies the possibility of freely ascribing intentions *post hoc* to past behavior in order for it to take on the shape of action. One possible objection is that non-observational first-person authority is needed for an agent with respect to her actions in order for her to be a rational agent in the first place. It seems that diachronic accounting consists in an external perspective on oneself and would thus exclude self-understanding or uphold possible self-estrangement or inauthenticity. Therefore, it seems that accepting diachronicity and weak agency leads to a constructivist self-interpretation where we unbind ourselves from our own past in accepting third personal theoretical descriptions of our behavior that we could not immediately and rationally endorse at the time of action. These descriptions are instead theoretically construed by taking an external perspective on ourselves.

Richard Moran thus argues that the therapeutic relation could contaminate first-personal authority:

> *In various familiar therapeutic contexts, for instance, the manner in which the analysand becomes aware of various of her beliefs and other attitudes does not necessarily conform to the Transparency Condition. The person who feels anger at the dead parent for having abandoned her, or who feels betrayed or deprived of something by another child, may only know of this attitude through the eliciting and interpreting of evidence of various kinds. She might become thoroughly convinced both from the constructions of the analyst, as well as from her own appreciation of the evidence, that this attitude must indeed be attributed to her. And yet, at the same time, when she reflects on the world-directed question itself, whether she has indeed been betrayed by this person, she may find that the answer is no or can't be settled one way or the other. So, transparency fails because she cannot learn of this attitude of hers by reflection on the object of that attitude. She can only learn of it in a fully theoretical manner, taking an empirical stance toward herself as a particular psychological subject." (Moran, 2001, p. 85).*

According to Moran, the difference between a theoretically formed perspective on oneself and a practical endorsement of one's attitude toward oneself remains even when therapy seem to have unearthed a historical truth for a person:

> *The person might be* told *of her feelings of betrayal, and she may not doubt this. But without her capacity to endorse or withhold endorsement from that attitude, and without the exercise of that capacity making a difference to what she feels, this information may as well be about some other person or about the voices in her head.*

> *From within a purely attributional awareness of herself, she is no more in a position to* speak for *her feelings than she was before, for she admits no authority over them. It is because her awareness of her sense of betrayal is detached from her sense of the reasons, if any, supporting it that she cannot become aware of it by reflecting on that very person, the one by whom she feels betrayed. The rationality of her response requires that she be in a position to* avow *her attitude toward him, and not just describe or report on it [.] (ibid., p. 93).*

For Moran, two features are central for self-understanding and thus for rational agency, namely, immediacy and transparency. Immediacy refers to the non-observational status of my self-knowledge. I immediately, without observation, evidence or inference know what my beliefs are, what I am doing, what I think. That is, immediacy is the epistemically privileged position I have toward my own mental states compared to the access of others to my mental states; they have to rely on observation and evidence-based reports to know something about my mental states (ibid., p. 92). For Moran, transparency is a condition on self-knowledge of beliefs that obtains when I determine which beliefs I have by reflecting on the worldly matter they concern. Transparency fails when the way I determine my beliefs is by reflecting on my own mental states or by observing my own behavior. Avowal of one's beliefs occurs when, as a result, of reflecting on the state of the world with which my beliefs are concerned, I come to endorse the beliefs in question.

We can see how transparency would fail in the case of weak agency and thus how the possibility for self-understanding would be excluded.[6] Moran describes the failure of transparency in the following way:

> *[F]or our analysand, if she is unable to learn of her attitude toward the person by whom she feels betrayed by thinking about* him, *if here she can only attribute beliefs to herself but cannot avow them, then she will not come to avow them by engaging in more and better attributions to herself. (The theoretical stance toward oneself constitutes itself as self-sufficient realm.) When I deliberate about something, the conclusion of my deliberation settles the question for me only in virtue of my attitude toward this activity, not in virtue of what I may belief about its effect on me. The aim and conclusion is the binding of oneself to a certain course of action (or proposition), not the production of a state of mind that I might then treat as (further) empirical evidence of how I should proceed." (Ibid., p. 95).*

Avowal, according to Moran, is the attitude with which I endorse the beliefs I have and when I make my first-person reports without any reference to evidence or inference. It is how I make up my mind and decide which beliefs of mine I endorse as true. The ability to avow my beliefs is constitutive my behavior as a rational agent: "A belief that cannot be avowed is thus cognitively isolated, unavailable to the normal processes of review and revision that constitute the rational health of belief and other attitudes. Thus we could explain why it is that the capacity not just for awareness of one's belief, but specifically awareness through avowal, is both the normal condition and part of the rational well-being of a person" (Ibid., p. 108). The difference between making up one's mind and "gesturing one's mind" (ibid., p.122), as Moran characterizes the analysand's verbal reports, is that in the first case, the agent responsibly and actively endorse her beliefs. Lying on the couch, verbally gesturing one's mind, one is passively dissociated from one's beliefs; one observes, finds evidence in one's own reports, or discovers in the reports—together with the analyst—something about oneself (ibid., p. 114 ff.) To both the analyst and oneself, such reports occur as data, as "more (verbal) behavior for interpretation" (ibid., p. 121). The agent does not speak her mind, she reports something she does not first personally avow, and her reports are treated as data, as indications of something else both by the analysand and the analyst, according to Moran.

Now, the point we find in Merleau-Ponty's quote above that the therapeutic relation consists in an existential commitment does not imply that we take on an empirical stance toward ourselves. Rather, the point is that, as part of an ongoing process of self-understanding, non-trivially, we come to discover certain things about ourselves that we then attempt to appropriate as actions of ours. If we accept Moran's account of self-knowledge, an objection to the notion of weak agency would be that any such agent would not have self-knowledge, as transparency would fail in the cases of weak agency since the agent would self-observe in order to speak and know her mind. To be clear, my use of the term self-understanding differs from Moran's notion of self-knowledge in the following way. In the broadest sense, both self-understanding and self-knowledge refer to what it means to know something about oneself. In the narrow sense, self-knowledge refers to the immediate and transparent way in which I avow my beliefs. For Moran, self-knowledge involved the ability to "avow one's state of mind and not merely to attribute it to oneself" (Moran, 2001, p. 100). It is tied to the Transparency Condition, according to which "a statement is made by consideration of the facts about X itself, and not by either an 'inward glance' or by observation of one's own behavior" (ibid., p. 101). This is Moran's technical sense of the term. In between the broad and the narrow sense, I use the term self-understanding to refer to the kind of understanding I have of myself over time, that is, when I come to realize something about myself. Whereas self-understanding can also be immediate, as when I find out that I have a stomach ache, it can also be diachronic as when I find out that I have been disengaged in a friendship over many years. The discovery I make does not exclude avowal; I can come to realize that I must end a friendship that I have been ending indirectly, and I take responsibility for my past intentional actions involved in being disengaged. As a disengaged friend, I did not immediately understand what I was doing, but diachronically, I appropriate the past doings as mine, and my self-understanding avows my former weak agency. My use of the term self-understanding is thus broader than Moran's technical use of the term self-knowledge, which only applies in cases that also involve immediacy and transparency. It is broader in that involves the temporal aspects of appropriation and thus of non-synchronic avowal; self-understanding includes diachronicity. In this particular sense, the notion of self-understanding that I apply resonate with some concerns tied specifically to the temporal aspect of Moran's notion of self-knowledge (see Lear, 2004; Webber, 2017).

[6] A comment on the difference between my use of the term self-understanding and Moran's use of the term self-knowledge will follow below.

In accepting something like weak agency, the claim is that avowal is possible over time and that one can learn to speak one's mind *diachronically*. In cases of weak agency, the dichotomy of relating either actively or passively to one's attitudes (Moran, 2001, p. 114) is not accepted; intentional action can be appropriated over time. I come to realize something about myself that I was not able to endorse or understand earlier, and thus, I take responsibility for something I did in the past, i.e., something that was not just an unreflective, passive behavior of mine. This is not a trivial reading of what coming to terms with being who one is means, nor is it a blatant antirealism that claims that, by diachronic appropriation, any past behavior could be turned into an action of mine. Rather, it means that we get to keep our initial action theoretic assumptions; only the first assumption must be reformulated as follows:

> *(1*) My behavior deserves the status of action, when the description under which my action is intentional is synchronically or diachronically accessible to me.*

In combination with the second assumption, we get the following: If I am responsible, then the description under which my action is intentional is accessible to me synchronically or diachronically. With this reformulation, it is emphasized that the possibility for appropriation is the requirement for something to be an action. It does not imply that everything can be turned into something for which I am responsible or that anything can be turned into an action of mine over time.

In finding oneself repeating certain patterns of behavior, my behavior can occur to me as questionable. The process of questioning can turn into a commitment not to repeat oneself. However, such a commitment is vulnerable and might have to be repeated itself. It also means that certain attempts at appropriation might lead to self-misunderstanding and thus will have to be revised, as they cannot be appropriated or endorsed after all [Freud (1938) is especially clear on this point]. Appropriation thus means that my current self-understanding is an ongoing existential commitment:

> *I take hold of my past and I transform it, I alter its sense, I free myself and I detach myself from it. But I only do so by committing myself elsewhere. Psychoanalytic treatment does not heal by provoking an insight into the past, but by first relating the subject to his doctor through new existential relations. It is not a question of giving a scientific approval to the psychoanalytic interpretation, nor of discovering a notional sense of the past; rather it is a question of re-living the past as signifying this or that, and the patient only achieves this by seeing his past from the perspective of his co-existence with the doctor. (Merleau-Ponty, 2012, p. 482).*

This movement or this strategy importantly involves a second-person perspective of engagement, not a third-person theoretical stance toward oneself. One commits oneself to the future by transforming what one already was into what one is through a responsive exchange of questioning one's past behavior. In this way, my past is not a set of fixed reasons, but some of aspects of my past behavior are open for questioning and thus for attempts at appropriation of past doings of mine in the form of intentional action. The self-understanding involved is not inauthentic in that the agent blindly accepts a third personal view about her past behavior; rather, it is a responsive exchange that leads one to realize something about oneself. Psychoanalysis is used by Merleau-Ponty to show that we are not radically free and that we come to understand ourselves through a process of commitment realized in the second person perspective, through an "I–Thou" relation that facilitates our relationship to the past that we are and that we can come to endorse diachronically. The therapeutic relation thus serves as an example of diachronic appropriation, as it displays the mode of questioning and existential commitment.

CONCLUDING REMARKS

The paper began by showing that certain cases are problematic for a standard theory of action, namely, cases of unreflective action where I intuitively feel responsible. I argued that the two basic and intuitive assumptions concerning the interrelation between consciousness, action, and responsibility can be kept if we accept a diachronic perspective on responsibility in cases of unreflective actions. I showed that cases of unreflective actions are not just in the periphery of the field of actions but are in fact at stake in culturally restricted patterns of behavior as well. Iris Young's notion of inhibited intentionality provided a case that allowed me to expand the scope of unreflective actions to embodied cultural life forms. I further argued that such cases are better understood as cases of weak agency. With the notion of weak agency, we can see how unreflective actions and habitual behavior are forms of action, as they can be diachronically appropriated. I used Merlau-Ponty's example of self-understanding in psychoanalysis in order to show how in appropriating our past behavior as action, we do not just take on any third personal description of our past that might seem suitable, but we assume responsibility for our past in coming to terms with being who we are. Appropriation is thus a rediscovery of one's reason for having acted in a certain way, and at the same time, appropriation is a process where one commits oneself to being and becoming who one is. Taking up a present by committing myself elsewhere is a different way of saying that I appropriate my past behavior anew by taking responsibility for who I am as a person. Finally, I proposed that self-understanding of this kind differs from immediate and transparent self-knowledge at stake in first-person endorsement of my actions and beliefs. I proposed a kind of self-understanding to be possible even in cases of weak agency; thus, I emphasized how the process of reason finding in diachronic appropriation is exemplified by the second-person responsive process of committing oneself to transformation as in the therapeutic context. The challenges for this position are many, but the main insight is that by reflecting on responsibility in cases of weak agency, we get a new approach to studying the role of consciousness as well as the possibility for self-understanding in unreflective actions.

DATA AVAILABILITY STATEMENT

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

AUTHOR CONTRIBUTIONS

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

FUNDING

The paper is part of the research project "Habitualization and the genesis of intention—Phenomenological analyses of transitions between action and behavior," which is funded by The Fritz Thyssen Foundation. In addition, this publication was supported by the Open Access Fund, Julius-Maximilians-Universität Würzburg, and by the Commission of Equality, the Faculty of Humanities and the Institute of Philosophy of Julius-Maximilians-Universität Würzburg.

ACKNOWLEDGMENTS

The author wishes to thank two reviewers for their very helpful, generous, and inspiring comments.

REFERENCES

Anscombe, G. E. M. (2001). *Intention*. Cambridge: Harvard University Press.

Bratman, M. (1987). *Intentions, Plans and Practical Reason*. Cambridge: Harvard University Press.

Crowell, S. (2013). *Normativity and Phenomenology in Husserl and Heidegger*. New York, NY: Cambridge University Press.

Davidson, D. (2001). *Agency in Essays on Actions and Events. 43-61*. Oxford: Clarendon Press.

Freud, S. (1909). "Notes on a case of obsessional neurosis," in *Standard Edition of the Complete Psychological Works of Sigmund Freud, (1956-1974)*, ed. J. E. Strachey (London: Hogarth Press).

Freud, S. (1914). "Remembering, repeating and working through," in *Standard Edition of the Complete Psychological Works of Sigmund Freud, (1956-1974)*, ed. J. E. Strachey (London: Hogarth Press).

Freud, S. (1938). Constructions in analysis. *Int. J. Psycho-Anal.* 19, 377–387.

Fuchs, T. (2012). "Body memory and the unconscious," in *Founding Psychoanalysis Phenomenologically*, eds D. Lohmar and J. Brudzinska (Dordrecht: Springer), 69–82. doi: 10.1007/978-94-007-1848-7_4

Fuchs, T. (2018). "Collective body memory," in *Embodiment, Enaction and Culture*, eds C. Durt, T. Fuchs, and C. Tewes (Masachusetts: MIT Press), 333–352.

Hampshire, S. (1970). *Thought and Action*. London: Chatto and Windus.

Honneth, A. (2008). *Reification: a New Look at an Old Idea*. Oxford: Oxford University Press.

Hursthouse, R. (1991). Arational actions. *J. Phil.* 88, 57–68. doi: 10.2307/2026906

Husserl, E. (1989). "Ideas pertaining to a Pure Phenomenology and to a phenomenological philosophy,"in *Second Book: Studies in the Phenomenology of Constitution*, eds R. Rojecewicz, and A. Schuwer (Dordrecht: Kluwer Academic Publishers).

Ingerslev, L. R. (2013). My body as an object: self-distance and social experience. *Phenomenol. Cogn. Sci.* 12, 163–178. doi: 10.1007/s11097-011-9228-9

Ingerslev, L. R. (2020). On the role of habit for self-understanding. *Phenomenol. Cogn. Sci.* 19, 481–497. doi: 10.1007/s11097-018-9605-8

Korsgaard, C. (1996). *The Sources of Normativity*. New York, NY: Cambridge University Press.

Lear, J. (1998). *Open Minded. Working Out the Logic of the Soul*. Cambridge MA: Harvard University Press.

Lear, J. (2004). Avowal and unfreedom. *Phil. Phenomenol. Res.* 69, 484–454.

Lear, J. (2017). *Wisdom Won From Illness. Essays in Philosophy and Psychoanalysis*. Cambridge, MA: Harvard University Press.

Merleau-Ponty, M. (2006). *The Structure of Behavior*, (A. Fischer, Trans). (Pittsburg: Duquesne University Press).

Merleau-Ponty, M. (2012). *Phenomenology of Perception*, (D. A. Landes, Trans). (London: Routledge).

Moran, R. (2001). *Authority and Estrangement. An Essay on Self-Knowledge*. Princeton, NJ: Princeton University Press.

Owens, D. (2017). Habitual agency. *Phil. Explorat.* 20, 93–108. doi: 10.1080/13869795.2017.1356358

Pickard, H. (2011). Responsibility without blame: empathy and the effective treatment of personality disorders. *Phil. Psychiatry Psychol.* 18, 209–244. doi: 10.1353/ppp.2011.0032

Pickard, H. (2015). Psychopathology and the ability to do otherwise. *Phil. Phenomenol. Res.* 90, 135–163. doi: 10.1111/phpr.12025

Setiya, K. (2017). *Practical Knowledge*. Oxford: Oxford University Press.

Steward, H. (2009). "Sub-intentional actions and the over-mentalization of agency," in *New Essays on the Explanation of Action*, ed. C. Sandis (London: Palgrave Macmillan).

Talero, M. (2006). Merleau-Ponty and the bodily subject of learning. *Int. Phil. Q.* 46, 191–203. doi: 10.5840/ipq20064622

Webber, J. (2017). "Habituation and first-person authority," in *Time and the Philosophy of Action*, ed. R. Altshuler and M. Sigrist (London: Routledge), 189–204.

Young, I. M. (1980). Throwing like a girl: a phenomenology of feminine body comportment, motility and spatiality. *Hum. Stud.* 3, 137–156. doi: 10.1007/bf02331805

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

ORIGINAL RESEARCH
published: 08 October 2020
doi: 10.3389/fpsyg.2020.530152

A Map of Consciousness Studies: Questions and Approaches

Takuya Niikawa[1,2*†]

[1] Institut Jean Nicod, ENS, Paris, France, [2] Faculty of Humanities and Human Sciences, Hokkaido University, Sapporo, Japan

This article aims to present a map of consciousness studies, which consists of a list of fundamental questions about consciousness and existing approaches to them. The question list includes five fundamental categories: Definitional, Phenomenological, Epistemological, Ontological, and Axiological. Each fundamental category is divided into more determinate questions. Existing approaches to each question are also classified into a few groups, presenting principal researchers who take each kind of approach. In the final section, I demonstrate the usefulness of the proposed map of consciousness studies by applying it to examine the integrated information theory and the global workspace theory of consciousness.

Keywords: consciousness, consciousness science, philosophy of consciousness, approach to consciousness, questions about consciousness

OPEN ACCESS

Edited by:
Christopher Gutland,
Sun Yat-sen University, China

Reviewed by:
Koji Ota,
Niigata University, Japan
Garrett Mindt,
Central European University, Hungary

****Correspondence:***
Takuya Niikawa
niitaku11@gmail.com

†Present address:
Takuya Niikawa,
Graduate School of Humanities, Kobe University, Kobe, Japan

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

Received: *28 January 2020*
Accepted: *02 September 2020*
Published: *08 October 2020*

Citation:
Niikawa T (2020) A Map of Consciousness Studies: Questions and Approaches. Front. Psychol. 11:530152. doi: 10.3389/fpsyg.2020.530152

INTRODUCTION

Academic research starts with research questions. An area of research typically develops by research questions being sophisticated, in particular, those being conceptually clarified and being divided into more determinate questions. In the philosophy of language, for instance, the research question of "what is the meaning of symbols?" was divided into two different types of questions, namely, the question about semantics—"what is the meaning of this or that symbol (for a particular person or group)?"—and the question about foundation—"in virtue of what facts about that person or group does the symbol have that meaning?" (Speaks, 2019, sec. 1). This division has helped us to develop theories of meaning without confusion. In linguistics, likewise, the research question of "what is the linguistic capacity?" can be divided into two distinct questions, namely, the competence question—"what is the linguistic competence?"—and the performance question—"what is the linguistic performance ability?" (Chomsky, 2014). This distinction helps us to develop theories of linguistic capacities while avoiding unnecessary confusion.

Consciousness studies have rapidly developed in the last three decades; many philosophical and scientific theories of consciousness have been proposed. However, it is far less clear how such theories of consciousness are related to each other. Some theories target different aspects of consciousness; some theories address the same aspect of consciousness but with different methodologies. Consider two influential scientific theories of consciousness, the integrated information theory (IIT) (Tononi, 2008; Tononi et al., 2016) and the global workspace theory (GWT) (Baars, 2005; Dehaene, 2014). Although many assume that they are competitive, it is unclear whether they are concerned with the same research subject in the first place (Ball, 2019). Given that there already exist many theories of consciousness, and it is far less clear how they are related, we need to stop trying to answer a specific research question set out in a theoretical framework for a moment and instead *take research questions about consciousness themselves as the target of investigation*. In other words, a *second-order investigation* of the research questions about consciousness is required to further develop consciousness studies.

As an initial step of the second-order investigation, this article presents *a systematic list of questions about consciousness* (see section "The List of Questions"). This list helps us to understand what questions the existing theories of consciousness address. In addition, the list helps each consciousness researcher to see what aspects of consciousness they are interested in.

After proposing a list of questions about consciousness, I also submit a list of *approaches* to each question (see section "The List of Approaches"). The list of approaches gives us the methodological overview of consciousness studies. It also helps researchers working in various fields to see what question they can tackle in their methodological/theoretical frameworks.

The list of questions is constructed by a *top–down approach*. I apply the traditional taxonomy of philosophical inquiries to categorize questions about consciousness. Thus, the proposed classificatory framework is neither arbitrary nor groundless. The list of approaches is constructed by a *bottom–up approach*. I take the existing approaches to each kind of question and then classify them based on their crucial methodological differences.

The final section of this article is dedicated to demonstrate the usefulness of the map of consciousness studies, which consists of the lists of questions and approaches, by applying it to examine IIT and GWT. I will argue that the proposed map is useful in that it can provide a multidimensional framework in which to compare various scientific theories of consciousness, including IIT and GWT.

THE LIST OF QUESTIONS

Philosophical inquiries have typically been divided into three categories: *Ontological*, *Epistemological*, and *Axiological* (Lee, 1966, p. 72; Woleński, 2004, p. 3). In addition to this traditional distinction, I incorporate two other fundamental categories into the classificatory framework for questions about consciousness, namely, *Definitional* and *Phenomenological*. Definitional inquiries explore satisfactory definitions of key concepts, such as "good" and "knowledge." The term "consciousness" is also a target of this inquiry. Phenomenology is a discipline in which to investigate conscious phenomena from the subjective point of view, which is typically distinguished from other disciplines of philosophy (Smith, 2018, sec. 1). There is no doubt that the category of phenomenology should be included in the classificatory framework for questions about consciousness.

Thus, we have five fundamental categories in which questions about consciousness are classified: *Definitional*, *Phenomenological*, *Epistemological*, *Ontological*, and *Axiological*[1]. Each fundamental category (except the definitional) has subcategories. The subcategories are set out partially in a bottom–up manner: it is partially based on the widely accepted division in the subject matter. In the rest of this section, I present the five fundamental questions about consciousness and how they are divided into subquestions (**Figure 1**).

[1]Note that I do not claim that the questions belonging to distinct fundamental categories are independent of each other. Rather, they are interrelated in such a way that the answer to one question affects the scope of possible answers to other kinds of questions.

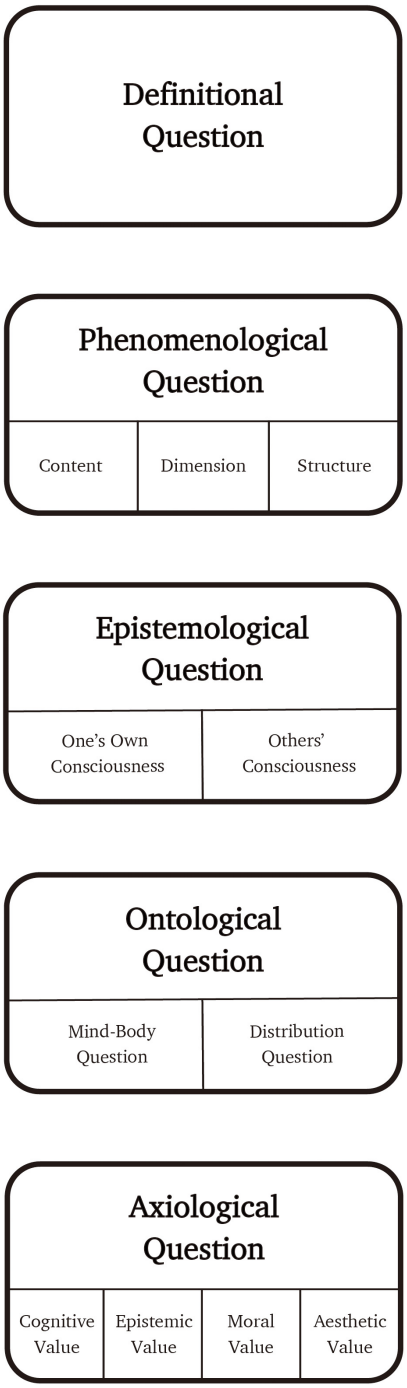


FIGURE 1 | The list of questions about consciousness.

Definitional Question

How should we define the term "consciousness" and its cognates?

The definitional question is not divided into subquestions. . . .

Phenomenological Question

What phenomenological features does consciousness have?

The phenomenological question is divided into three subquestions, depending on what aspect of consciousness to focus on, namely, *content*, *dimension*, or *structure*. The content of consciousness is understood as *variable* features of consciousness, such as experienced color, shape, movement, taste, or feel[2]. The dimension of consciousness is understood as the fundamentally different kinds of conscious experiences, such as perceptual, cognitive, and emotional dimensions (Kriegel, 2015). The structure of consciousness is understood as *invariable* features of consciousness, such as unity and figure-ground structure (Bayne, 2010; Watzl, 2011; Macpherson, 2015). While the general structures of consciousness itself are typically discussed in the philosophy of consciousness, the specific structures of each dimension of consciousness can also be investigated.

Content Question
What content does consciousness have?

Dimension Question
What dimensions does consciousness have?

Structure Question
What structures does consciousness have?

Epistemological Question
How do we know about consciousness?

The epistemological question is divided into two subquestions, depending on whose consciousness to address, whether *one's own consciousness* or *the consciousness of others*.

Epistemological Question About One's Own Consciousness
How do we know about our own consciousness?

Epistemological Question About Others' Consciousness
How do we know about the consciousness of others?

Ontological Question
How is consciousness located in the world?

The ontological question is divided into two subquestions[3]. The first concerns the relation between consciousness and the physical world; the second concerns the distribution of consciousness over the physical world.

Mind–Body Question
What relation holds between consciousness and the physical world (in particular our brain)?

Distribution Question
How is consciousness distributed in the physical world? (In other words, what has consciousness?)

Axiological Question
What values does consciousness have?

This question is divided into four subquestions, depending on what kind of value to address, namely, *cognitive*, *epistemic*, *moral*, or *aesthetic*[4].

Cognitive Value Question
What type of cognitive capacity does consciousness enable its possessor to have?

Epistemic Value Question
What type of knowledge does consciousness enable its possessor to have?

Moral Value Question
What type of moral status does consciousness enable its possessor to have?

Aesthetic Value Question
What type of aesthetic value does consciousness enable its possessor to have?

THE LIST OF APPROACHES

In this section, I present approaches to each kind of questions that have been actually employed by consciousness researchers with a brief assessment of them (**Figure 2**). Note that although each approach can be taken individually to address one question, we can also take different approaches in combination to address one specific question. In this sense, these approaches are not exclusive.

Approaches to the Definitional Question
Let us start with the definitional question: How should we define the term "consciousness" and its cognates? There are two approaches to the definitional question: (I) *example-based approach* and (II) *essence-based approach*.

The example-based approach defines the term "consciousness" as *something* that is shared by typical examples of conscious states/experiences, such as pain experience and visual experience (Velmans, 2009; Nida-Rümelin, 2016; Prinz, 2016; Schwitzgebel, 2016). This approach can provide a theoretically neutral definition of consciousness, since it does not refer to any distinctive property in the definition of the term "consciousness." The problem of the example-based approach is that it is unclear how we should determine the scope of typical examples of conscious experiences. If we restrict the "typical" examples

[2]For the overview of philosophical issues about the content of consciousness, see Macpherson (2011) and Siegel (2016).

[3]One might wonder why I do not formulate the ontological question as "what is consciousness?" The reason is that it is ambiguous in that it can also be interpreted as the definitional question and as the phenomenological question asking the essential phenomenological features of consciousness. This ambiguity of the question of "what is consciousness?" may cause confusions in consciousness studies.

[4]Although I owe this division partly to Kriegel (2019), there are two differences. First, I do not introduce the distinction between intrinsic and instrumental values for the sake of simplicity. Second, I add cognitive value because (1) it does not seem to be reducible to other kinds of value and (2) the cognitive value question has actually been discussed in scientific consciousness studies (Kanai et al., 2019).

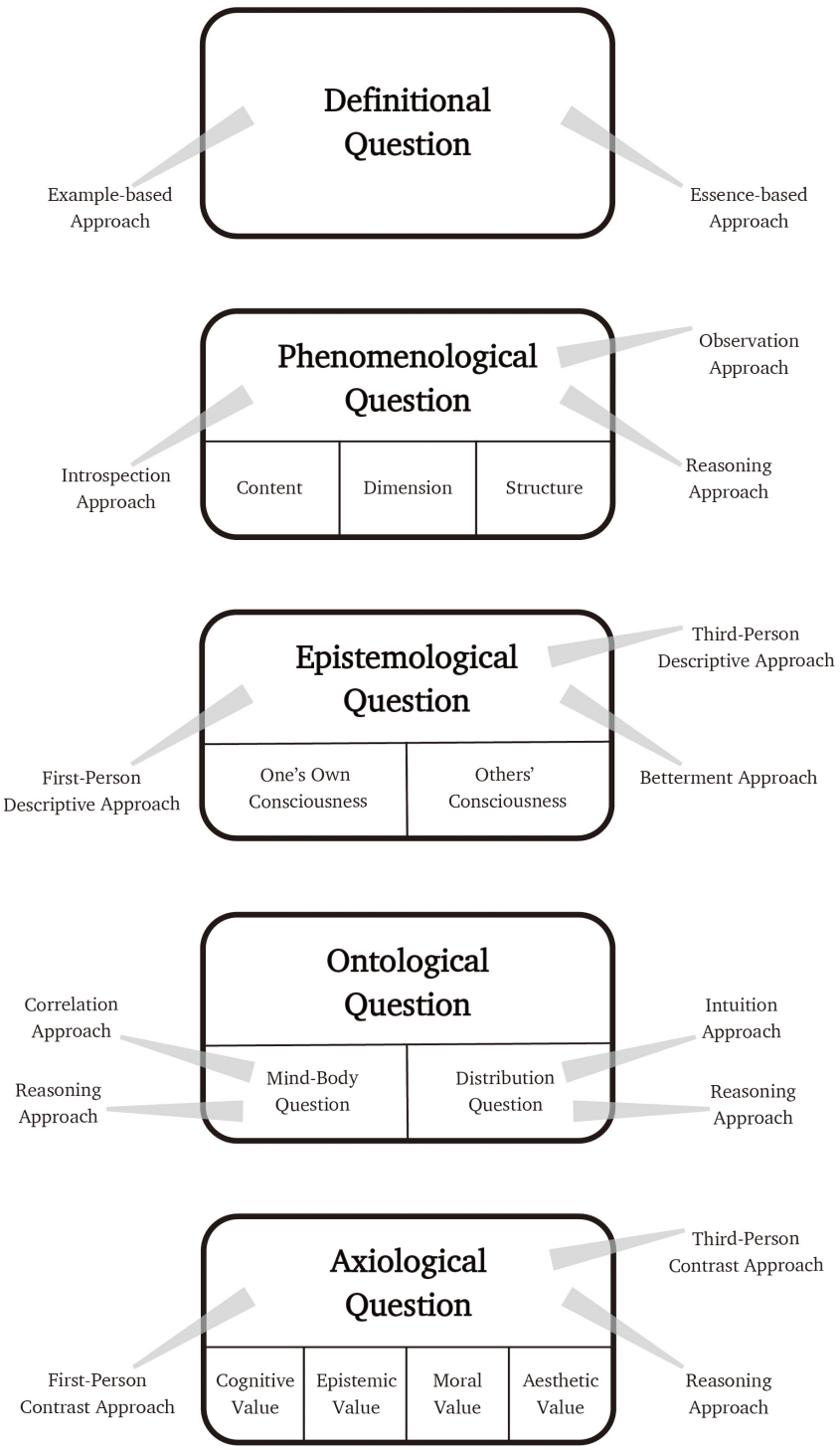


FIGURE 2 | The list of questions about consciousness and each approach to them.

of conscious experiences too narrowly, borderline cases of consciousness, such as dreamless sleep and vegetative states, may be automatically excluded from consciousness studies. However, it is controversial whether dreamless sleep and vegetative states are conscious states (Shea and Bayne, 2010; Windt et al., 2016).

The essence-based approach defines the term "consciousness" by referring to its essential property such as phenomenality (or "what-it-is-like-ness") (Chalmers, 1997), the property of being inner, qualitative, and subjective (Searle, 2000), and being accurately reportable (Baars, 1993, p. 19). The merit of this

approach is that it can provide an informative definition of the term "consciousness." Its problem is that it is highly controversial what property we should count as the definitional property of consciousness; there may be no single property that all the consciousness researchers accept to be essential for consciousness. Thus, the essence-based approach may cause a dispute over the correct definition of consciousness. This dispute can be resolved if we take a pluralist position about the definition of consciousness, allowing that there are several different notions of consciousness defined in different manners, such as "phenomenal consciousness" and "access consciousness" (Block, 1995). However, meta-level questions arise for the pluralists: what relation holds between those notions of consciousness? Are they different aspects of a single phenomenon, or do they refer to different phenomena?

Approaches to the Phenomenological Question

The second question is the phenomenological question: What phenomenological features does consciousness have? This question is divided into three subquestions: content, dimension, and structure. There are three approaches to each subquestion: (I) *introspection approach*, (II) *observation approach*, and (III) *reasoning approach*.

The introspection approach explores the phenomenological features of conscious experiences by *introspection in a broad sense*, where it involves not only the cognitive activities of "turning one's attention inward and attending to one's own concurrent internal goings-on" but also uses of memory, imagination, and concept application (Kriegel, 2015, pp. 20, 21). Thus, introspection in this sense can include imagining a conscious experience and conceptually describing its phenomenological feature; it can also involve imaginatively comparing a current conscious experience with past conscious experiences[5].

The introspection approach works for basic contents of conscious experiences and their salient dimensions and structures. For instance, when I introspect on my current conscious experience of drinking Springbank 15 years, I can find that it has sweetness, smokiness, saltiness, and its distinctive sulfur smell as its flavor contents, that it has temporal continuity as its structure, and that it has perceptual and algedonic dimensions. However, there are cases for which the introspection approach does not work. For instance, sophisticated skills of introspection are required to capture the complex phenomenological features of consciousness such as dynamical interactions of attentional shifts and flavor profiles of drinking Springbank 15 years, but people typically lack such skills. Furthermore, introspection seems unable to determine whether a perceptual experience can have causal relations (e.g., touching a screen causing the screen flashing) and natural kind properties (e.g., being water) as its content (Siegel, 2007) and whether consciousness has cognitive phenomenology as a distinct dimension (Spener, 2011). This is not because our skills of introspection are not sophisticated to the required extent, but because it is unclear how introspective data are related to these issues. We need to clarify what introspective data are predicted if we consciously experience causal relations between events in addition to experiencing sequential occurrences of events, if we consciously experience the property of being water in addition to experiencing the presence of clear, colorless liquid, and if consciousness has a cognitive dimension as being irreducible to other dimensions such as sensory and imaginative ones. Introspection turns out useless if no introspectable difference is predicted there. A more fundamental limitation of introspection is that one's introspection is not effective to understand the phenomenological features of conscious experiences that one is *unable to have*. For instance, it is hard for normal people to understand by introspection the phenomenological features of schizophrenic experiences and synesthetic experiences.

The observation approach is to infer what phenomenological features a conscious experience has from its possessor's observable states/behaviors, including their subjective reports. For instance, when one reports that she sees a red patch, then we can infer that her conscious experience has red-color content; when one groans painfully, we can infer that she has pain experience[6]. This approach is available to explore the phenomenological features of conscious experiences that are difficult for researchers to have by themselves, such as schizophrenic and synesthetic experiences. We can, in principle, infer what phenomenological features such conscious experiences have from their possessors' observable states/behaviors, in particular, their introspective reports (Hubbard and Ramachandran, 2003; Fuchs, 2007; Simmonds-Moore, 2016). Experimental neuroscience of consciousness typically employs the observation approach in investigating neural properties responsible for the phenomenological features of conscious experiences (Tononi and Koch, 2015, sec. 3). This is in part because it needs to collect quantitative data about neural activities from *many experimental participants* rather than the researchers themselves alone. The practical problem with this approach lies in how to interpret the observable states/behaviors. In addressing untypical conscious experiences, even the introspective reports of such experiences are difficult to interpret. Here is a report from a patient with depersonalization syndrome: "I feel as though I'm not alive as though my body is an empty, lifeless shell. I seem to be standing apart from the rest of the world, as though I'm not really here" (Bockner, 1949, p. 969). It is hard to understand what the patient's experience is like. Furthermore, Michel (2019) points to the crucial role of background beliefs in interpreting observable states/behaviors, claiming that the disagreements among consciousness scientists mainly lie in whether to interpret certain observable states/behaviors as evidence for the presence of a phenomenological feature.

The *reasoning approach* infers what phenomenological features consciousness has from a given thesis about

[5]Husserlian Phenomenology can be counted as introspective on this broad definition of introspection (Breyer and Gutland, 2016, p. 13; Gutland, 2018).

[6]It is controversial whether some kind of inference is always required to know what others consciously experience. This paper sets aside this issue for simplicity. See McDowell (1983) for an argument against the necessity of inference.

consciousness. For instance, Pitt (2004) argues that given that the possession of consciousness enables us to introspectively know what we think, consciousness must have a cognitive dimension as being irreducible to other dimensions[7]. A challenge to those who adopt this approach is to justify a thesis about consciousness used as a premise for reasoning. Note that the premise in Pitt's reasoning can be counted as an answer to the epistemic value question: What type of knowledge does consciousness enable its possessor to have? As shown in this, answers to other kinds of questions, in particular, the *ontological question* and the *axiological question*, can possibly be the premises used in the reasoning approach.

Approaches to the Epistemological Questions

The third question is the epistemological question: How do we know about consciousness? This question is divided into two subquestions depending on whose consciousness to address, namely, one's own consciousness or the consciousness of others. There are three approaches to the epistemological question: (I) *first-person descriptive approach*, (II) *third-person descriptive approach*, and (III) *betterment approach*.

The first-person descriptive approach describes how we know about our own consciousness and the consciousness of others from the first-person perspective. To describe how one knows about one's own consciousness from the first-person perspective is to describe first-personally *the processes of introspection*. This approach is employed in constructing/assessing theories of introspection. For example, some philosophers reflect on the process of introspection on perceptual experience and describe it as being "transparent" in that we know the contents of our own perceptual experience through being aware of the external objects/events (Harman, 1990; Tye, 2000). A theory of introspection is constructed/assessed partially based on the transparency of perceptual experience. If a theory of introspection implies that introspection is entirely distinct from perceptual awareness, the theory seems to conflict with the transparency of perceptual experience and therefore be assessed negatively. Likewise, to describe how one knows about the consciousness of others from the first-person perspective is to describe first-personally *the processes of knowing others' conscious experiences*. Some philosophical and phenomenological accounts of how to know others' conscious experiences are partially based on the first-personal descriptions of such processes (Wittgenstein, 1980, sec. 570; Scheler, 2008; Overgaard, 2017).

The third-person descriptive approach describes how we know about our own consciousness and the consciousness of others from the third-person perspective. To describe how one knows about one's own consciousness from the third-personal perspective is to describe third-personally the processes of introspection. This approach typically focuses on the neural/psychological processes responsible for introspection (Fleming et al., 2010; Baird et al., 2013; Jacobs and Silvanto, 2015), where introspection is not differentiated from metacognition [for the conceptual relation between introspection and metacognition, see Overgaard and Sandberg (2012, sec. 1)]. A theory of introspection can also be evaluated based on the relevant scientific findings. Likewise, to describe how one knows about the consciousness of others from the third-personal perspective is to describe third-personally the processes of knowing others' conscious experiences. This approach includes attempts to describe the processes of *mind reading* and *empathy*. Some focus on the relevant neural/psychological processes (Marsh, 2018), others address external conditions in which we try to know others' conscious experiences (Gallagher and Hutto, 2008). A theory of mind reading/empathy can be constructed/assessed based on the data acquired through this type of exploration.

The betterment approach explores how we can better know about consciousness, rather than just describing how we know about it. There are a few research projects that can be counted as the betterment approach. Some training programs to enhance the skills of empathy have been developed (Lam et al., 2011; Englander, 2014). Likewise, there is a research project to design a training program to enhance the skills of introspection in general (Miyahara et al., 2020). There are also many attempts to invent an interview-based method to know better what others consciously experience (Petitmengin, 2006; Langdridge, 2007; Giorgi, 2009; Petitmengin et al., 2019). Moreover, brain-decoding techniques may be available to know better about the consciousness of others, including behaviorally non-responsive patients' experiences (Naci et al., 2017). Importantly, we can see the betterment approach as developing methods to address the phenomenological question[8].

Approaches to the Ontological Question

The fourth question is the ontological question: How is consciousness located in the world? This question is divided into two subquestions. The first is the mind–body question: What relation holds between consciousness and the physical world (in particular our brain?) The second is the distribution question: How is consciousness distributed in the physical world? (In other words, what has consciousness?) There are two approaches to the mind–body question: (I) *correlation approach* and (II) *reasoning approach*. There are also two approaches to the distribution question: (I) *intuition approach* and (II) *reasoning approach*.

The correlation approach explores what neural or informational feature is correlated with the presence of a phenomenological feature of consciousness (or the presence of consciousness itself) by using brain scanning technologies such as functional magnetic resonance imaging (fMRI) and brain stimulation techniques such as repetitive transcranial magnetic stimulation (rTMS). Simply put, this is to explore "neural correlates of consciousness (NCC)," which are the minimal neuronal mechanisms jointly sufficient for a specific

[7]This is not the original form of Pitt's argument but its possible reformulation.

[8]One might wonder why the question of "how it is possible for us to know about consciousness," which has been much discus-sed in philosophy, is not included in the list of subcategories of the epistemological question. The reason is that the "how possible" question should be interpreted as a form of the ontological question, namely: what relation must hold between consciousness and the world (including our body and brain) given that it is possible for us to know about consciousness?

content, dimension, or structure of consciousness (or the presence of consciousness itself) (Crick and Koch, 1990). There are many findings of the neural and informational correlates of consciousness (Koch et al., 2016; Boly et al., 2017; Wu, 2018, secs 4, 5). For example, some found that the conscious experience of a visual scene is correlated with the activities of the parahippocampal place area of our brain (a subregion of the parahippocampal cortex that lies medially in the inferior temporo-occipital cortex) (Mégevand et al., 2014); others found that the conscious experience of a human face is correlated with the activities of the posterior and mid fusiform gyrus (Parvizi et al., 2012). The limitation of the correlation approach is that it cannot, in principle, reveal a more substantial relationship between consciousness and the physical world than the correlation relation. Since the correlation relation is consistent with many metaphysical relations such as causal relation, grounding relation, and identity relation, the correlation approach cannot determine which metaphysical relation holds between consciousness and the physical world (Kozuch and Kriegel, 2015).

The reasoning approach infers what relation holds between consciousness and the physical world from a given thesis about consciousness. For example, Papineau (2002, pp. 31–35) takes the causal efficacy thesis that consciousness can cause physical effects as a key premise for reasoning and argues that consciousness is identical to physical properties. Chalmers (2010, pp. 106–108) takes the conceivability of a phenomenal zombie—the thesis that it is conceivable that a physical duplicate of us lacks consciousness—as a key premise for reasoning and argues that consciousness cannot be physical. Campbell (2002, chap. 6) argues that perceptual consciousness must be constituted by ordinary mind-independent objects on the premise that perceptual consciousness enables its possessor to know about such ordinary mind-independent objects demonstratively. As shown in these examples, the reasoning approach can address what metaphysical relation holds between consciousness and the physical world beyond mere correlation. As we saw in section "Approaches to the Phenomenological Question", however, a challenge to those who adopt this approach is to justify the thesis about consciousness used as a key premise for reasoning. To address this, for example, one may try to justify the causal efficacy thesis by appealing to our folk psychological briefs, such as the one that "my conscious thirst caused me to fetch a beer" (Papineau, 2002, p. 21); another may try to justify the conceivability of a phenomenal zombie by providing an argument against the *a priori* entailment from physical facts to phenomenal facts. The essential difficulty with the reasoning approach is to settle the conflicts between those who take distinct theses, which are justified in different manners, as the premises to infer opposing ontological positions (such as physicalism and anti-physicalism).

The intuition approach to the distribution question asks our intuition what has consciousness. We typically have some intuitive thoughts about what can have consciousness. For instance, it seems doubtless to me that other human beings are conscious. Many other kinds of mammals, such as dogs and cats, seem to have consciousness. However, microphysical entities and machines like my laptop do not seem to have consciousness. It is unclear to me whether insects and plants are conscious. On the assumption that intuition is a reliable epistemic route to know about the distribution of consciousness over the world, we can employ our intuition to answer the distribution question. The obvious problem with this approach is to justify the assumption that intuition is reliable with respect to the distribution of consciousness over the world.

The reasoning approach is also available to address the distribution question: to infer what has consciousness from a given thesis about consciousness. For example, if we take biological naturalism that consciousness is a biological phenomenon (Searle, 1992) as a premise for reasoning, we can infer that non-biological entities, such as machines and robots, cannot have consciousness. If we take IIT that consciousness is identical to internally generated and integrated information (Tononi, 2008) as a premise for reasoning, we can conclude that any system that generates information in an integrated manner has consciousness (for the detail of IIT, see section "Applications: Integrated Information Theory and Global Workspace Theory"). As we have seen, the essential difficulty with this approach is to settle the debates between those who take distinct theses, which are justified in different manners, as the premises for reasoning.

Note that the answer to the distribution question directly affects the scope of the phenomenological and epistemological questions. For instance, since IIT implies that computers which generate information in an integrated manner possess consciousness, the question of how we can know about the consciousness of such computers arises for advocates of IIT. Likewise, IIT opens up the phenomenological question about such computers: What content, dimension, and structure does their consciousness have?

Approaches to the Axiological Question

The fifth question is the axiological question: What values does consciousness have? This question is divided into four subquestions depending on what kind of value to address: cognitive, epistemic, moral, and esthetic. The scope of those subquestions is not restricted to the values of *consciousness itself* but includes those of each content, dimension, and structure of consciousness. There are three approaches to the axiological question: (I) *first-person contrast approach, (II) third-person contrast approach,* and (III) *reasoning approach*.

The first-person contrast approach explores what difference there is in relevant value between the cases where one has and lacks consciousness (or where one's consciousness has and lacks a specific phenomenological feature) from *the first-person perspective*. This approach typically consists of the following two steps: (a) to first-personally imagine that one loses consciousness (or a specific phenomenological feature disappears from one's consciousness) and (b) to consider what value-related feature she thereby loses. Siewert (1998, 2014) takes this approach, arguing that (1) consciousness makes the life of its possessor worth living and that (2) only the possessor of consciousness can perform intentional cognitive activities/processes such as making judgments and having desires. The first point concerns moral value and possibly aesthetic value; the second point

concerns cognitive value. Campbell (2002, chap. 1) also takes this approach, arguing that perceptual experience enables its possessor to know about ordinary mind-independent objects demonstratively. The problem with this approach is that it is controversial whether our first-personal thoughts about the values of consciousness are reliable. When reflecting on how we visually discriminate an object from others, for example, we are likely to think that if we lose perceptual consciousness, we cannot carry out the discrimination task. However, this intuitive thought seems to be falsified by the case of blindsight: the patient can achieve various visual discrimination tasks by "guesswork," even though he said he did not have any visual experience (Weiskrantz, 2007).

The third-person contrast approach explores what difference there is in relevant value between the cases where one has and lacks consciousness (or where one's consciousness has and lacks a specific phenomenological feature) from *the third-person perspective*. Dehaene and Naccache (2001) take this approach, arguing that consciousness enables durable and explicit information maintenance, novel combinations of operations, and intentional behavior. Weiskrantz (1997, chap. 7) focuses on a broad spectrum of syndromes in which there seems to be a loss of capacities related to consciousness, such as blindsight and aphasic disorders, arguing that consciousness grounds the capacities to perform flexible thinking and imagining. Kriegel (2017) compares our natural attitudes to conscious beings and non-conscious beings, and argues that consciousness confers dignity as a moral status on its possessors (possibly with some other conditions). However, it may be objected that the apparent differences in values can be explained without appealing to consciousness (Lau, 2009; see also Rosenthal, 2008). Hence, a challenge to the third-person contrast approach is to argue that the proposed difference in a type of value cannot be well explained without referring to consciousness.

The reasoning approach infers what values consciousness has from a given thesis about consciousness. For example, Tye (1996) takes the representationalist thesis that consciousness is representational as the key premise for reasoning and concludes that consciousness enables its possessor "to do a wide variety of things that they would not be able to do without it—for example, to recognize objects, to avoid knocking into them" (pp. 301, 302). The proponents of the attentional schema thesis that consciousness is an internal model of attention (Graziano and Webb, 2014) can take it as a premise for reasoning and conclude that consciousness enables its possessor to control attention in proper manners.

APPLICATIONS: INTEGRATED INFORMATION THEORY AND GLOBAL WORKSPACE THEORY

This article proposes a map of consciousness studies, which consists of a systematic list of questions about consciousness and existing approaches to each question. In this final section, I apply this map to examine IIT and GWT. I first address how IIT answers each fundamental question that I have listed. In doing so, I point out several challenges to IIT. I then take the same procedure to examine GWT. I finally propose a way to clarify the relation between IIT and GWT with the help of the proposed map of consciousness studies. The discussion is sketchy but still sufficient to demonstrate how the proposed map can be used to examine and compare theories of consciousness.

Let us start with the definitional question. Tononi (2015, abstract, emphasis added) claims that IIT "attempts to identify the *essential properties* of consciousness (*axioms*) and, from there, infers the properties of physical systems that can account for it (postulates)." He lists five essential properties of consciousness, namely, intrinsic existence, composition, information, integration, and exclusion, and calls them "axioms" (Tononi, 2015, sec. 2). The intrinsic existence axiom states that consciousness exists independently from external observers, the composition axiom states that consciousness is structured, the information axiom states that each conscious experience is the particular way it is and thereby it differs from other possible conscious experiences, the integration axiom states that consciousness is unified, and the exclusion axiom states that consciousness is definite in content and spatiotemporal grain[9]. The fact that they are called "axioms" suggests that the conjunction of the listed essential properties fixes the *reference* of "consciousness." Thus, IIT takes the essence-based approach to the definitional question, claiming that consciousness is defined in terms of the five axioms.

One slogan of IIT is that it goes "from phenomenology to physics" (Tononi et al., 2016, p. 450); the axioms are called the "phenomenological axioms" (Oizumi et al., 2014). This indicates that the axioms are derived from phenomenological considerations, namely, by addressing the phenomenological question, in particular, the structure question of what invariant features consciousness has (since the essential properties of consciousness are the invariant of consciousness). This suggests that advocates of IIT answer the definitional question through tackling the structure question.

Advocates of IIT claim that the phenomenological axioms "cannot be doubted and do not need proof" and are "taken to be immediately evident" (Oizumi et al., 2014, p. 2). This shows that they take the introspection approach to the structure question, rather than the observation approach and the reasoning approach, to derive the phenomenological axioms. However, some philosophers cast doubt on the plausibility of the axioms as capturing the essential phenomenological features of consciousness (Bayne, 2018; Pokropski, 2018; Miyahara and Witkowski, 2019). This demonstrates that the phenomenological axioms *can be doubted* and should not be taken to be immediately evident. Thus, advocates of IIT must justify the phenomenological axioms, employing the other approaches if needed.

Let us next move onto the ontological question. IIT specifies five informational features of physical systems (so-called "postulates"), each of which is supposed to account

[9]For more details of axioms in IIT, see Oizumi et al. (2014); Tononi et al. (2016), and Miyahara and Witkowski (2019).

for a corresponding phenomenological axiom, and states that every physical system that realizes the five postulates possesses consciousness[10]. This statement is counted as the answer to the distribution question. Nevertheless, it is not fully clear what reasoning is in play here (especially in what sense each postulate "accounts for" a corresponding phenomenological axiom and why each postulate necessitates the phenomenological feature represented by the corresponding axiom). In order to evaluate IIT's answer to the distribution question, thus, we should clarify the exact premises and inferential steps that constitute the reasoning in question.

IIT answers the mind–body question by stating that conscious experience is identical to an integrated informational structure of physical systems that instantiates the five postulates (Tononi, 2015, sec. 4). There is, however, no mention of how the identity claim is derived in any IIT literature. As we have seen in section "Approaches to the Ontological Question", identity is not reasonably inferred only from the presence of correlation, since other metaphysical relations such as causal relation and grounding relation are also compatible with the presence of correlation. To justify the identity claim, advocates of IIT need to clarify what theses they use as the premises for the reasoning in question, in addition to the experimental finding that there is a correlation between the presence of consciousness and a relevant informational structure of brains (Massimini et al., 2005). Otherwise, we cannot properly evaluate IIT's identity claim.

Let us finally examine what implications IIT have for the epistemological and the axiological questions. First, IIT seems to have an implication for the epistemological question about the consciousness of others. IIT states that the phenomenological features of consciousness (in particular contents and dimensions) are reflected in the *form* of the informational structure of physical systems (Tononi, 2015, sec. 4; Tononi et al., 2016, p. 459). It follows from this that we can infer the phenomenological features of the consciousness of others from the form of the informational structure of their brain, which we can, in principle, specify from the third-person perspective. This can be counted as an answer to the epistemological question about the consciousness of others. IIT also has an implication for the cognitive value question. If it is cognitively advantageous for physical systems to generate information in an integrated manner, IIT implies that the possession of consciousness is cognitively advantageous for that very reason.

I turn to how GWT (in particular its major advocate Stanislas Dehaene) answers each fundamental question listed in The List of Questions section. Dehaene (2014, pp. 8, 9) defines consciousness in terms of "conscious access": the content of mental state/process is consciously accessible if and only if it enters awareness and becomes reportable to others. This definition consists of two notions, *awareness* and *reportability*. The property of being reportable serves to provide an informative definition of consciousness, since we can set out an objective procedure to determine whether a piece of information is reportable for its possessor. In contrast, it is unclear how "awareness" is different from "consciousness" in our ordinary conceptual understandings. Furthermore, it is unclear what behavioral standard can be used to determine whether one is aware of a piece of information, *as being different from the one for reportability*. Nevertheless, Dehaene does not seem to provide an analytic explanation of the notion of awareness. Instead, he presents a few examples of *being aware of something*. For instance, he presents an example of visual illusion and states:

> *Twelve dots, printed in light gray, surround a black cross. Now stare intently at the central cross. After a few seconds, you should see some of the gray dots fade in and out of existence. For a few seconds, they vanish from your* awareness*; then they pop back in. Sometimes the entire set goes away, temporarily leaving you with a blank page—only to return a few seconds later with a seemingly darker shade of gray. (Dehaene, 2014, p. 17)*

This suggests that Dehaene leads his readers to grasp the sense of "awareness" through the examples presented in his book. If this is correct, his definition of consciousness is not entirely operational, for it does not reduce the sense of "consciousness" to reportability alone. In defining consciousness, Dehaene seems to take the example-based and essence-based approaches in combination; the former corresponds to the "awareness" part, and the latter corresponds to the "reportability" part.

Dehaene (2014, chap. 4) takes the correlation approach to the mind–body question, presenting many relevant empirical findings[11]. Based on them, he identifies four physiological markers that index whether a stimulus is consciously accessible:

> *First, a conscious stimulus causes an intense neuronal activation that leads to a sudden ignition of parietal and prefrontal circuits. Second, in the EEG, conscious access is accompanied by a slow wave called the P3 wave, which emerges as late as one-third of a second after the stimulus. Third, conscious ignition also triggers a late and sudden burst of high-frequency oscillations. Finally, many regions exchange bidirectional and synchronized messages over long distances in the cortex, thus forming a global brain web. (Dehaene, 2014, pp. 158, 159)*

Dehaene then provides a functionalist account as to why consciousness is correlated with those physiological makers.

> *The human brain has developed efficient long-distance networks, particularly in the prefrontal cortex, to select relevant information and disseminate it throughout the brain. Consciousness is an evolved device that allows us to attend to a piece of information and keep it active within this broadcasting system. Once the information is conscious, it can be flexibly routed to other areas according to our current goals. Thus we can name it, evaluate it, memorize it, or use it to plan the future (Dehaene, 2014, p. 161).*

This functionalist account describes how a piece of information is cognitively processed in our brain when it is consciously accessible and thereby explains why the above physiological makers occur in functional terms. This account is thus an empirically supported correlation-based answer to the mind–body question. This is, I think, the core thesis of GWT. However, Dehaene (2014, p. 161) goes beyond the

[10]For the details of the postulates, see Oizumi et al. (2014).

[11]It is necessary to address the content question before exploring the correlation between the contents of consciousness and neural/informational features of brains. Dehaene (2014, p. 41–45) takes the observation approach to address the content question, emphasizing the availability of subjective reports from experimental participants as objective data.

empirically supported claim regarding correlation, claiming that "consciousness is brain-wide information sharing." If we interpreted this statement literally, it would mean the identity between consciousness and the brain-wide information sharing. However, this identity claim does not directly follow from the empirically supported claim about correlation. If Dehaene (2014) defined consciousness only in terms of reportability, then the identity claim would be derived from the fact that reportability can be reductively explained in terms of brain-wide information sharing. However, Dehaene (2014) includes "awareness" in his definition of consciousness, which is supposed to be grasped through examples. It is unclear whether *the property of being aware of something* is considered to be reductively explained in functional terms, unlike reportability. Thus, Dehaene is required to explain why *the property of being aware of something* should be counted as standing in the identity relation, rather than other metaphysical relations, to the brain-wide information sharing. As in the case of IIT, we cannot properly evaluate IIT's identity claim unless some explanation is provided.

Global workspace theory has implications for (i) the cognitive value question, (ii) the epistemological question about the consciousness of others, and (iii) the distribution question. Given that a piece of information can be flexibly routed to many brain areas only when it is consciously accessible, it is plausible to think that (i) consciousness enables its possessor to process information in such flexible manners (Dehaene, 2014, chap. 3) and that (ii) we can know about the content of the consciousness of others by detecting the information widely shared in his/her brain. (iii) It follows from GWT's identity claim that every creature who has "brain-wide information sharing" is conscious (Dehaene, 2014, chap. 6.7).

We can clarify the relation between IIT and GWT by comparing their answers to each fundamental question. Let us take three questions, for example, the definitional question, the mind–body question, and the distribution question. For the definitional question, IIT states that consciousness is defined in terms of the five phenomenological axioms, which are supposed to capture the essential properties of consciousness. In contrast, GWT defines consciousness in terms of awareness and reportability. By comparing the two definitions of consciousness, we can examine *whether IIT and GWT have the same research subject in the first place*. For the mind–body question, IIT states that conscious experience is identical to an integrated informational structure of physical systems that instantiates the five postulates. In contrast, GWT states that consciousness is brain-wide information sharing. By comparing the two identity claims, we can examine *whether they are compatible or conflicting*. For the distribution question, IIT states that every physical system that realizes the five postulates possesses consciousness. In contrast, GWT implies that every creature who has brain-wide information sharing has consciousness. By examining whether each kind of creature overlaps, we can see *whether IIT and GWT substantially differ in what existing creatures/entities have consciousness*. In this way, we can conduct a *multidimensional comparison* between IIT and GWT. This enables us to assess the two theories systematically and comparatively in the multidimensional evaluative space.

I close this article by presenting three ideas on how to proceed with consciousness research with the help of the lists of questions and approaches proposed in this article. First, we should examine how existing theories of consciousness answer each fundamental question about consciousness and what approach their advocates adopt. By doing so, we can obtain systematic understandings of each theory of consciousness, which enable us to see what part of each theory of consciousness needs to be justified and developed. Second, we should conduct a multidimensional comparison of the existing theories of consciousness. This enables us to obtain a detailed and well-organized review of how they are related to each other. These two points have been demonstrated in the discussions of IIT and GWT. Third, each research group should clarify what question and approach to take in investigating consciousness. By doing so, they can be aware of the scope, limitation, and potential implications of their research project and also of its relations to existing theories of consciousness.

Although I believe that the proposed lists of questions and approaches contribute to the development of consciousness studies, I do not think that they are entirely satisfactory. The map of consciousness studies presented in this article can be revised and further enriched. I hope that this article also works as a springboard for a further second-order investigation on consciousness studies as being distinct from the first-order investigation on consciousness.

DATA AVAILABILITY STATEMENT

The original contributions presented in the study are included in the article/supplementary material.

AUTHOR CONTRIBUTIONS

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

FUNDING

I gratefully acknowledge the support of the Japan Society for the Promotion of Science (JSPS KAKENHI Grant Number: 18K00031).

ACKNOWLEDGMENTS

I would like to thank Benedicte Veillet, Ying-Tung Lin, Katsunori Miyahara, Matthieu Koroma, Lok-Chi Chan, and Uriah Kriegel for their detailed comments on this article. I also thank the audience at the 22nd conference of the Association for the Scientific Study of Consciousness, the 10th Annual Meeting of Japan Association for Contemporary and Applied Philosophy, the consciousness workshop at National Yang-Ming University, and Paris Consciousness/Self-consciousness Group for their questions and comments. I appreciate Nanami Yamada for designing the figures in this article.

REFERENCES

Baars, B. J. (1993). *A Cognitive Theory of Consciousness. Reprint edition*. Cambridge: Cambridge University Press.

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

Baird, B., Smallwood, J., Gorgolewski, K. J., and Margulies, D. S. (2013). Medial and lateral networks in anterior prefrontal cortex support metacognitive ability for memory and perception. *J. Neurosci.* 33, 16657–16665. doi: 10.1523/JNEUROSCI.0786-13.2013

Ball, P. (2019). *Neuroscience Readies for a Showdown Over Consciousness Ideas*. New York, NY: Quanta Magazine.

Bayne, T. (2010). *The Unity of Consciousness*. Oxford: Oxford University Press.

Bayne, T. (2018). On the axiomatic foundations of the integrated information theory of consciousness. *Neurosci. Conscious.* 4:niy007. doi: 10.1093/nc/niy007

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

Bockner, S. (1949). The depersonalization syndrome: report of a case. *J. Ment. Sci.* 95, 968–971. doi: 10.1192/bjp.95.401.968

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

Breyer, T., and Gutland, C. (2016). *Phenomenology of Thinking: Philosophical Investigations into the Character of Cognitive Experiences*. London: Routledge. doi: 10.1523/jneurosci.3218-16.2017

Campbell, J. (2002). *Reference and Consciousness*. Oxford: Oxford University Press.

Chalmers, D. (1997). *The Conscious Mind: In Search of a Fundamental Theory. Revised* 版. New York, NY: Oxford University Press.

Chalmers, D. (2010). *The Character of Consciousness*. Oxford: Oxford University Press.

Chomsky, N. (2014). *Aspects of the Theory of Syntax*. Anniversary, Reprint 版. Cambridge, MA: The MIT Press.

Crick, F., and Koch, C. (1990). Towards a neurobiological theory of consciousness. *Semin. Neurosci.* 2, 263–275.

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

Dehaene, S., and Naccache, L. (2001). Towards a cognitive neuroscience of consciousness: basic evidence and a workspace framework. *Cognition* 79, 1–37. doi: 10.1016/S0010-0277(00)00123-2

Englander, M. (2014). Empathy training from a phenomenological perspective. *J. Phenomenol. Psychol.* 45, 5–26. doi: 10.1163/15691624-12341266

Fleming, S. M., Weil, R. S., Nagy, Z., Dolan, R. J., and Rees, G. (2010). Relating introspective accuracy to individual differences in brain structure. *Science* 329, 1541–1543. doi: 10.1126/science.1191883

Fuchs, T. (2007). The temporal structure of intentionality and its disturbance in schizophrenia. *PSP* 40, 229–235. doi: 10.1159/000101365

Gallagher, S., and Hutto, D. D. (2008). "Understanding others through primary interaction and narrative practice," in *The Shared Mind: Perspectives on Intersubjectivity Converging Evidence in Language and Communication Research (CELCR)*, eds T. P. Racine, J. Zlatev, C. Sinha and E. Itkonen (Amsterdam: John Benjamins Publishing Company), 17–38. doi: 10.1075/celcr.12.04gal

Giorgi, A. (2009). *The Descriptive Phenomenological Method in Psychology: A modIfied Husserlian Approach*. Pittsburgh, PA: Duquesne University Press.

Graziano, M. S. A., and Webb, T. W. (2014). A mechanistic theory of consciousness. *Int. J. Mach. Conscious.* 06, 163–176. doi: 10.1142/S1793843014400174

Gutland, C. (2018). Husserlian phenomenology as a kind of introspection. *Front. Psychol.* 9:896. doi: 10.3389/fpsyg.2018.00896

Harman, G. (1990). The intrinsic quality of experience. *Philos. Perspect.* 4, 31–52. doi: 10.2307/2214186

Hubbard, E., and Ramachandran, V. (2003). The phenomenology of synaesthesia. *J. Conscious. Stud.* 10, 49–57.

Jacobs, C., and Silvanto, J. (2015). How is working memory content consciously experienced? The 'conscious copy' model of WM introspection. *Neurosci. Biobehav. Rev.* 55, 510–519. doi: 10.1016/j.neubiorev.2015.06.003

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

Koch, C., Massimini, M., Boly, M., and Tononi, G. (2016). Neural correlates of consciousness: progress and problems. *Nat. Rev. Neurosci.* 17, 307–321. doi: 10.1038/nrn.2016.22

Kozuch, B., and Kriegel, U. (2015). "Correlation, causation, constitution: on the interplay between the science and philosophy of consciousness," in *The Constitution of Phenomenal Consciousness*, ed. S. M. Miller (Amsterdam: John Benjamins), 400–417. doi: 10.1075/aicr.92.17koz

Kriegel, U. (2015). *The Varieties of Consciousness*. Oxford: Oxford University Press.

Kriegel, U. (2017). "Dignity and the phenomenology of recognition-respect," in *Emotional Experiences: Ethical and Social Significance*, eds J. J. Drummond and S. Rinofner-Kreidl (New York, NY: Rowman & Littlefield), 121–136.

Kriegel, U. (2019). The value of consciousness. *Analysis* 79, 503–520. doi: 10.1093/analys/anz045

Lam, T. C. M., Kolomitro, K., and Alamparambil, F. C. (2011). Empathy training: methods. Evaluation practices, and validity. *J. MultiDiscipl. Eval.* 7, 162–200.

Langdridge, D. (2007). *Phenomenological Psychology: Theory, Research and Method*. Harlow: Pearson Education.

Lau, H. (2009). Volition and the function of consciousness. *Faith Philos.* 26, 537–552. doi: 10.5840/faithphil200926554

Lee, D. S. (1966). "Ultimacy and the philosophic field of metaphysics," in *Metaphysics and Belief Tulane Studies in Philosophy*, ed. F. R. Merlan (Dordrecht: Springer Netherlands), 71–102. doi: 10.1007/978-94-010-3540-8_5

Macpherson, F. (2011). "Introduction: the admissible contents of experience," in *The Admissible Contents of Experience*, eds K. Hawley and F. Macpherson (Hoboken, NJ: John Wiley & Sons, Ltd), 1–15. doi: 10.1002/9781444343915.ch1

Macpherson, F. (2015). the structure of experience, the nature of the visual, and type 2 blindsight. *Conscious. Cogn.* 32, 104–128. doi: 10.1016/j.concog.2014.10.011

Marsh, A. A. (2018). The neuroscience of empathy. *Curr. Opin. Behav. Sci.* 19, 110–115. doi: 10.1016/j.cobeha.2017.12.016

Massimini, M., Ferrarelli, F., Huber, R., Esser, S. K., Singh, H., and Tononi, G. (2005). Breakdown of cortical effective connectivity during sleep. *Science* 309, 2228–2232. doi: 10.1126/science.1117256

McDowell, J. (1983). "Criteria, defeasibility, and knowledge," in *Proceedings of the British Academy 68*, Oxford: Oxford University Press, 455–479.

Mégevand, P., Groppe, D. M., Goldfinger, M. S., Hwang, S. T., Kingsley, P. B., Davidesco, I., et al. (2014). Seeing scenes: topographic visual hallucinations evoked by direct electrical stimulation of the parahippocampal place area. *J. Neurosci.* 34, 5399–5405. doi: 10.1523/JNEUROSCI.5202-13.2014

Michel, M. (2019). Consciousness science underdetermined. *Ergo Open Access J. Philos.* 6, 2019–2020. doi: 10.3998/ergo.12405314.0006.028

Miyahara, K., Niikawa, T., Hamada, H. T., and Nishida, S. (2020). Developing a short-term phenomenological training program: a report of methodological lessons. *New Ideas Psychol.* 58:100780. doi: 10.1016/j.newideapsych.2020.100780

Miyahara, K., and Witkowski, O. (2019). The integrated structure of consciousness: phenomenal content, subjective attitude, and noetic complex. *Phenomenol. Cogn. Sci.* 18, 731–758. doi: 10.1007/s11097-018-9608-5

Naci, L., Sinai, L., and Owen, A. M. (2017). Detecting and interpreting conscious experiences in behaviorally non-responsive patients. *NeuroImage* 145, 304–313. doi: 10.1016/j.neuroimage.2015.11.059

Nida-Rümelin, M. (2016). "The illusion of illusionism," in *Illusionism as a Theory of Consciousness*, ed. K. Frankish (England: Imprint Academic), 200–214.

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

Overgaard, M., and Sandberg, K. (2012). Kinds of access: different methods for report reveal different kinds of metacognitive access. *Philos. Trans. R. Soc. Lond. B Biol. Sci.* 367, 1287–1296. doi: 10.1098/rstb.2011.0425

Overgaard, S. (2017). Other minds embodied. *Cont. Philos. Rev.* 50, 65–80. doi: 10.1007/s11007-016-9388-y

Papineau, D. (2002). *Thinking About Consciousness*. Oxford: Oxford University Press.

Parvizi, J., Jacques, C., Foster, B. L., Withoft, N., Rangarajan, V., Weiner, K. S., et al. (2012). Electrical stimulation of human fusiform face-selective regions distorts face perception. *J. Neurosci.* 32, 14915–14920. doi: 10.1523/JNEUROSCI.2609-12.2012

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

Petitmengin, C., Remillieux, A., and Valenzuela-Moguillansky, C. (2019). Discovering the structures of lived experience. *Phenomenol. Cogn. Sci.* 18, 691–730. doi: 10.1007/s11097-018-9597-4

Pitt, D. (2004). The phenomenology of cognition or "What Is It like to Think That P?". *Philos. Phenomenol. Res.* 69, 1–36. doi: 10.1111/j.1933-1592.2004.tb00382.x

Pokropski, M. (2018). Commentary: from the phenomenology to the mechanisms of consciousness: integrated information theory 3.0. *Front. Psychol.* 9:101. doi: 10.3389/fpsyg.2018.00101

Prinz, J. (2016). Against illusionism. *J. Conscious. Stud.* 23, 186–196.

Rosenthal, D. M. (2008). Consciousness and its function. *Neuropsychologia* 46, 829–840. doi: 10.1016/j.neuropsychologia.2007.11.012

Scheler, M. (2008). *The Nature of Sympathy*. Piscataway, NJ: AldineTransaction.

Schwitzgebel, E. (2016).). Phenomenal consciousness, defined and defended as innocently as i can manage. *J. Conscious. Stud.* 23, 224–235.

Searle, J. R. (1992). *The Rediscovery of the Mind*. Cambridge, MA: Bradford Books.

Searle, J. R. (2000). Consciousness. *Annu. Rev. Neurosci.* 23, 557–578. doi: 10.1146/annurev.neuro.23.1.557

Shea, N., and Bayne, T. (2010). The vegetative state and the science of consciousness. *Br. J. Philos. Sci.* 61, 459–484. doi: 10.1093/bjps/axp046

Siegel, S. (2007). How can we discover the contents of experience? *Southern J. Philos.* 45, 127–142. doi: 10.1111/j.2041-6962.2007.tb00118.x

Siegel, S. (2016). "The contents of perception," in *The Stanford Encyclopedia of Philosophy*, ed. E. N. Zalta (Stanford, CA: Stanford University).

Siewert, C. (1998). *The Significance of Consciousness*. Princeton, NJ: Princeton University Press.

Siewert, C. (2014). "Speaking up for consciousness," in *Current Controversies in Philosophy of Mind*, ed. U. Kriegel (London: Routledge), 199–221. doi: 10.4324/9780203116623-9

Simmonds-Moore, C. A. (2016). An interpretative phenomenological analysis exploring synesthesia as an exceptional experience: insights for consciousness and cognition. *Q. Res. Psychol.* 13, 303–327. doi: 10.1080/14780887.2016.1205693

Smith, D. W. (2018). "Phenomenology," in *The Stanford Encyclopedia of Philosophy*, ed. E. N. Zalta (Stanford, CA: Stanford University).

Speaks, J. (2019). "Theories of meaning," in *The Stanford Encyclopedia of Philosophy*, ed. E. N. Zalta (Stanford, CA: Stanford University).

Spener, M. (2011). "Disagreement about cognitive phenomenology," in *Cognitive Phenomenology*, eds T. Bayne and M. Montague (Oxford: Oxford University Press), 268–284. doi: 10.1093/acprof:oso/9780199579938.003.0012

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

Tononi, G. (2015). Integrated information theory. *Scholarpedia* 10:4164. doi: 10.4249/scholarpedia.4164

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

Tononi, G., and Koch, C. (2015). Consciousness: here, there and everywhere? *Phil. Trans. R. Soc. B* 370:20140167. doi: 10.1098/rstb.2014.0167

Tye, M. (1996). The function of consciousness. *Noûs* 30, 287–305. doi: 10.2307/2216271

Tye, M. (2000). *Consciousness, Color, and Content*. Cambridge, MA: MIT Press.

Velmans, P. M. (2009). How to define consciousness—and how not to define Consciousness. *J. Conscious. Stud.* 16, 139–156.

Watzl, S. (2011). "Attention as structuring of the stream of consciousness," in *Attention: Philosophical and Psychological Essays*, eds C. Mole, D. Smithies, and W. Wu (Oxford: Oxford University Press), 145–173.

Weiskrantz, L. (1997). *Consciousness Lost and Found: A Neuropsychological Exploration*. New York, NY: Oxford University Press.

Weiskrantz, L. (2007). Blindsight. *Scholarpedia* 2:3047. doi: 10.4249/scholarpedia.3047

Windt, J. M., Nielsen, T., and Thompson, E. (2016). Does consciousness disappear in dreamless sleep? *Trends Cogn. Sci.* 20, 871–882. doi: 10.1016/j.tics.2016.09.006

Wittgenstein, L. (1980). *Remarks on the Philosophy of Psychology*. Oxford: Blackwell.

Woleński, J. (2004). "The history of epistemology," in *Handbook of Epistemology*, eds I. Niiniluoto, M. Sintonen, and J. Woleński (Dordrecht: Springer Netherlands), 3–54. doi: 10.1007/978-1-4020-1986-9_1

Wu, W. (2018). "The neuroscience of consciousness," in *The Stanford Encyclopedia of Philosophy*, ed. E. N. Zalta (Stanford, CA: Stanford University).

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

METHODS
published: 03 March 2021
doi: 10.3389/fpsyg.2021.602081

Framing a Phenomenological Mixed Method: From Inspiration to Guidance

Kristian Moltke Martiny[1,2]*, *Juan Toro*[1,2] *and Simon Høffding*[3,4]

[1] Center for Subjectivity Research, Faculty of Humanities, University of Copenhagen, Copenhagen, Denmark, [2] The Enactlab, Copenhagen, Denmark, [3] Department of Sports Science and Biomechanics, University of Southern Denmark, Odense, Denmark, [4] RITMO Centre for Interdisciplinary Studies in Rhythm, Time, and Motion, University of Oslo, Oslo, Norway

Despite a long history of researchers who combine phenomenology with qualitative or quantitative methods, there are only few examples of working with a phenomenological mixed method—a method where phenomenology informs both qualitative and quantitative data generation, analysis, and interpretation. Researchers have argued that in working with a phenomenological mixed method, there should be mutual constraint and enlightenment between the qualitative (first-person, subjective) and quantitative (third-person, objective) methods for studying consciousness. In this article, we discuss what a framework for phenomenological mixed methods could look like and we aim to provide guidance of how to work within such framework. We are inspired by resources coming from research in mixed methods and existing examples of phenomenological mixed-method research. We also present three cases of phenomenological mixed methods where we study complex social phenomena and discuss the process of how we conducted the studies. From both the research inspiration and our own studies, we depict the landscape of possibilities available for those interested in mixing phenomenology with qualitative and quantitative methods, as well as the challenges and common pitfalls that researchers face. To navigate in this landscape, we develop a three-fold structure, focusing on (1) the phenomenological frame, (2) the phenomenologically informed generation of qualitative and quantitative data (tier one), and (3) the phenomenologically informed analysis and interpretation of data (tier two).

Keywords: phenomenology, mixed method, phenomenological interview, musical communication and absorption, research in cerebral palsy

OPEN ACCESS

Edited by:
Wenjing Cai,
Shanghai Jiao Tong University, China

Reviewed by:
Anthony Chemero,
University of Cincinnati, United States
Philipp Schmidt,
Julius Maximilian University of Würzburg, Germany

***Correspondence:**
Kristian Moltke Martiny
kmartiny@hum.ku.dk

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

Received: *02 September 2020*
Accepted: *20 January 2021*
Published: *03 March 2021*

Citation:
Martiny KM, Toro J and Høffding S (2021) Framing a Phenomenological Mixed Method: From Inspiration to Guidance. Front. Psychol. 12:602081. doi: 10.3389/fpsyg.2021.602081

INTRODUCTION

How do we investigate consciousness with its manifold nuances and complexities? Philosophers working within the philosophical tradition of phenomenology have, since its inception, tried to answer this methodological question, while breaking up disciplinary frontiers and working in interdisciplinary contexts. Despite this effort, there are fewer examples of working with a phenomenological mixed method. "Mixing" is here used as an umbrella term to refer to the multifaceted procedures of combining, integrating, linking, and employing multiple methods (Creswell, 2003; Creswell et al., 2003). Following Tashakkori et al. work, we define mixed-method investigations as "research in which the investigator collects and analyses data, integrates the

findings and draws inferences using both qualitative and quantitative approaches" (Tashakkori and Creswell, 2007, 3). To work with a phenomenological mixed method is consequently to phenomenologically inform both the qualitative and quantitative data generation, analysis, and interpretation.

One of the few examples of a phenomenological mixed method is called "Neurophenomenology" as developed by Varela in relation to the proposal of naturalizing phenomenology (Varela, 1996; Varela and Shear, 1999). Varela argued that in order for cognitive science to work as a scientific method for studying consciousness, a mutual constraint [also called mutual enlightenment (Gallagher, 1997)] should exist between first-person (qualitative) and third-person (quantitative) methods in generating, analyzing, and validating both subjective and objective data. Philosophical phenomenology was used as the theoretical foundation for framing the mutual constraint and enlightening the two methods.

Guidelines of how to work with such neurophenomenological mixed methods have been developed under the heading of micro-phenomenology (Bitbol and Petitmengin, 2017). As we describe and discuss below, this way of working focuses on investigating the microdynamic processes and pre-reflective aspects of experience. Finding inspiration in micro- and neurophenomenology, we believe it is still necessary to show a variety of different ways to work with phenomenological mixed methods to be able to understand different kinds of experiences at different pre-reflective and reflective levels.

Our aim is to depict a landscape of possibilities available for those interested in this way of working. In this article, we will therefore discuss more broadly what a framework for phenomenological mixed methods could look like. In the attempt to develop a framework, we engaged ourselves in three studies to understand the process of working with phenomenological mixed methods.

The first study concerns the phenomenon of joint musical absorption. Having collaborated with the Danish String Quartet to establish definitions and criteria for various forms of joint musical absorption (Høffding 2019), we were interested to see if and how bio-rhythms were implicated herein. Would a strong sense of joint, "interkinesthetic absorption" (Høffding, 2019, 233-40) be matched by synchronized breathing and heart rate, or would physiology and phenomenology be divorced in this situation? This was the main research question driving our first mixed-method study. The methodological challenge we therefore had to solve was how we could investigate, both qualitatively and quantitatively, the experiences of musicians, while they were playing.

In addition to this study, we engaged in two mixed-method studies on physical disability. The first study focused on understanding joint actions involving people with cerebral palsy (CP), which is a disorder that occurs due to a non-progressive lesion in the developing central nervous system, damaging sensorimotor predictive models and processes and causing limitations to bodily functionality (i.e., bodily coordination and adjustments). In embodied and enactive accounts of cooperation which are phenomenologically informed, it is argued that there exists a strong and direct link between coordination in joint actions and the interactors' positive experiences of connectedness, harmony and flow (Fuchs and De Jaegher, 2009; Marsh et al., 2009; Fantasia et al., 2014). This means that people who have problems with bodily coordination, such as people with CP, would be expected to experience these positive feelings to a lesser degree, or not at all, making their experience of joint actions negative. Our research question was therefore: How do people with CP, who have problems with bodily coordination, experience positive and negative joint actions? To answer this question, we needed to design a mixed-method study where we could relate bodily coordination and positive or negative experiences for joint actions in the case of CP. However, this raised the methodological challenge of how to create positive and negative joint actions in a way where we could also investigate their functional and affective aspects?

The second study on physical disability focused on reducing prejudice toward people with physical disability. Most research on prejudice applies the Contact hypothesis (Binder et al., 2009). Its basic assumption is that contact—under appropriate conditions—will reduce prejudice and negative attitudes between in- and out-groups and between majority and minority group members—which in our cases means between people with and without physical disability. Inspired by the phenomenologically informed "interactive turn" in social cognitive science (e.g., de Jaegher et al., 2010; De Bruin et al., 2012; Schilbach et al., 2013; Satne and Roepstorff, 2015), our hypothesis was that embodied engagement as contact would reduce prejudice toward physical disability by changing the attitude to be more positive toward people with physical disabilities. To investigate this hypothesis, we needed a mixed-method design where we could test peoples' explicit (conscious) and implicit (unconscious) processes of attitude formation before and after an embodiment and engagement-based intervention. The methodological challenge we were dealing with was therefore: How do we change experiences of prejudice, while at the same time investigate the process of change?

In order to answer to these three challenges and provide a frame for phenomenological mixed methods, we need to review and develop a number of perspectives and analyses. In the next section Inspiration: Mixing With Phenomenology, we discuss what we mean by "phenomenological" in a mixed-method context, look at challenges found in the literature for working with phenomenological mixed methods, and discuss two inspirational examples, namely, microphenomenology (Bitbol and Petitmengin, 2017) and the EASE interview (Parnas et al., 2005) from phenomenological psychopathology. This inspiration was the point of departure from which we engaged in our three studies and which helped us in dealing with the methodological challenges, which we present in section Process: Three Cases of How to Use Phenomenological Mixed Methods.

In section Guidance: Steps, Decisions, and Standards, we combine the inspiration from the previous work with the lessons learned in conducting our studies. We develop and clarify a three-part structure that can serve as an overall guideline for the phenomenological mixed-method researcher. The structure consists of:

The phenomenological frame: The philosophical foundation, commitments, theories, concepts, and distinctions that frame two intricately linked tiers—starting from the theoretical point of departure and continuing through to the answers developed in response to the research question.
Tier one: The phenomenologically informed qualitative and quantitative data generation.
Tier two: The phenomenologically informed qualitative and quantitative data analysis and interpretation.

Although the three parts are presented as a chronological series of first, second, and third steps to take in order to conduct a phenomenological mixed method, the three parts are intricately linked and conducting phenomenological mixed-method research requires many steps back and forth between the three parts. We therefore do not believe that a procedure-like, step-by-step manual can be developed for researchers to follow within each of the three parts. As will be clear in the article, there is not one paradigmatic way to work with a phenomenological mixed method, but different ways in which qualitative and quantitative methods can be mixed within a phenomenological frame. Nevertheless, the aim of the three-part structure is to provide guidance and help those interested in phenomenological mixed methods take steps and make decisions that are performative and phenomenologically consistent.

INSPIRATION: MIXING WITH PHENOMENOLOGY

To understand what we mean by a phenomenological mixed method, it is useful to start by defining what is meant by "phenomenological" in a mixed method context.

Phenomenological Foundations and Commitments

Fundamentally, the aim of phenomenology is to attain an understanding and description of the structures of human experience. It aims at being a rigorous science of consciousness, by pursuing "the things themselves" and taking experience seriously. This means letting the descriptions of conscious experience themselves come to the fore, withholding pre-established theories, explanations, and beliefs about the objects of conscious experience. It also means that phenomenology is opposed to the belief in the metaphysical realism that fuels various objectivist, scientistic, and naturalistic approaches (Gallagher and Zahavi, 2008). Metaphysical realism defends a mind- and experience-independent reality that can only truly and objectively be discovered by reducing and/or eliminating aspects of human experience and subjectivity. Such belief is, according to phenomenology, based on the objectivist illusion that our experience of worldly objects is irrelevant (or rather, an obstacle) when it comes to determining how objects really are, thus denying subjectivity any foundational ontological function, and further that it is possible to obtain a pure, absolute, and objective perspective on reality (Zahavi, 2017). In contrast, phenomenology argues that any understanding of the world comes from the first-person perspective of someone, even if this someone is a scientist who aims to take up an objective and third-person perspective in her research.

With such criticism of objectivism and naturalism, one might think that phenomenology would be better poised to include methods coming from qualitative research rather than quantitative research. One might further question whether it is at all possible to use phenomenology to mix both types of methods. In mixed-method research, this is called the problem of commensurability, namely, that mixing does not seem possible without contradiction, because the different methods reflect epistemologies and ontologies that are not compatible (Small, 2011).

In mixed-method research, scholars have not pointed to phenomenology as a solution to the problem of commensurability, but rather turned to pragmatism as their theoretical foundation (e.g., Rallis and Rossman, 2003; Johnson and Onwuegbuzie, 2004; Greene, 2007; Morgan, 2007; Denscombe, 2008). The basic principle in pragmatism for mixed methods is that the act of discovery should be prioritized over the theoretical justifications for knowledge. The researcher should apply whatever means and methods she finds useful for answering her research questions (Creswell and Plano Clark, 2011). However, in using phenomenology as a theoretical foundation for mixed methods, we neither need nor want to settle for a methodological strategy of "whatever works." In fact, working with phenomenology is a way of prioritizing both the act of discovery and its theoretical justification.

Phenomenology argues that any understanding of the world comes from the first-person perspective of someone, but this does not mean that the aim is to develop a subjective account of experience. As with any scientific approach, phenomenology strives to avoid arbitrary or biased accounts of experience that focus solely and particularly on idiosyncratic experiences. Instead, phenomenology focuses on idiosyncratic experiences in order to understand and describe their invariant structures. Further, it emphasizes the interdependence of subjectivity and objectivity:

> "the phenomenologists' focus on the first-person perspective is as much motivated by an attempt to understand the nature of objectivity, as by an interest in the subjectivity of consciousness. Indeed, rather than taking the objective world as the point of departure, phenomenology asks how something like objectivity is possible in the first place." (Gallagher and Zahavi, 2008, 24)

Unlike more pragmatic, relativistic, and postmodern positions, phenomenology, at least on Gallagher and Zahavi's interpretation, does not deny the existence of objectivity, nor is it antiscientific, even if it surely questions the meaning of both objectivity and science. So this position does, indeed, poise phenomenology as a good starting point to mix methods that pertain to experiential or subjective aspects as found in qualitative methods and so-called objective aspects as found in the natural sciences and quantitative methods. This is also evidenced by the recent discussion on "naturalizing phenomenology" (Varela, 1996; Petitot et al., 1999; Varela and Shear, 1999; Zahavi, 2004).

In Høffding and Martiny (2016), we attempted to provide a framework for phenomenologically grounded work with interviews in qualitative research. Using this framework as an inspiration, the four principle commitments for working with a phenomenological mixed method are as follows:

1. To the thing itself: Using qualitative and quantitative methods to acquire detailed first-person and third-person understanding of an experience in question.
2. Invariant structures: Using qualitative and quantitative methods to grasp the invariant structures of the experience.
3. Subjectivity cannot be reduced to objectivity: In working with the qualitative and quantitative methods, the first-person perspective needs to be understood on its own terms, rather than reducing it to objective descriptions or deducing the qualitative from the quantitative.
4. Enaction, embodiment, and embeddedness: Phenomenology construes subjectivity and objectivity as embodied, enactive, and embedded. Qualitative and quantitative methods directly confront us with these aspects of experience.

Applying these commitments, we wish to maintain the complexity and irreducibility of conscious experience and the interdependence or co-constitution of subjectivity and objectivity. As emphasized in principle 4, this is done by understanding the research process in phenomenological mixed method as a social practice that rests on the enacted and embodied observations, experiences, and expertise of the individual researchers, but which is developed into shared knowledge of a research community through an intersubjective sense-making process (Depraz et al., 2003; Gallagher and Zahavi, 2008; Martiny, 2017). In other words, the research process includes a group of researchers with different perspectives based in either first-person, qualitative or third-person, quantitative science. Both of these perspectives are embedded and contextualized by the second-person, intersubjective perspective of the community. The aim of phenomenological mixed methods is to meaningfully integrate these three perspectives in order to make sense of the data and (hopefully) answer the research question.

Front-Loading: Solving the Challenges of Ignorance and Hyperphilosophizing

Determining how to phenomenologically frame the research in the beginning of one's mixed-method research process is not without challenges. There is a spectrum between ignoring one's phenomenological point of departure and being overly and unproductively focused on it, in such a way that it leads to "hyperphilosophizing."

In a recent and intense debate about how phenomenology should inform qualitative research debate, Zahavi has both criticized qualitative researchers for belittling and ignoring the contributions of phenomenologists like Husserl, Heidegger, and Merleau-Ponty (Zahavi, 2019b), and for hyperphilosophizing and misinterpreting their contributions such as the phenomenological method of epoché and the reduction (Zahavi, 2019a). Based on his critique, Zahavi suggests a productive and pragmatic way for qualitative researchers to work with a phenomenological point of departure. They should familiarize themselves with the phenomenological theory and its philosophical origin but refrain from focusing on its orthodoxy and directly adopting Husserl's, Heidegger's, or Merleau-Ponty's philosophical method by including methodological steps that are irrelevant for qualitative researchers such as epoché and the phenomenological reduction. They should therefore be informed by the comprehensive theoretical framework (ideas, concepts, and distinctions) that philosophical phenomenology has to offer, so that it makes sense in a qualitative research context and allows for better qualitative research results.

Zahavi and Martiny (2019) provide examples of how to pragmatically apply phenomenology in qualitative research. These examples—two of which we discuss below—can also be used as inspiration for phenomenological mixed methods. At the core of these examples, and one solution to the challenges of ignorance and hyperphilosophizing, is the idea of applying phenomenology by using phenomenological concepts, distinctions and theory, rather than its method.

One way to do this is to use the method of "front-loading," which Gallagher (2003) originally proposed in order to work phenomenologically with experiments in cognitive science, and Køster and Fernandez (in press) recently proposed in order to phenomenologically ground qualitative research. More specifically, Gallagher writes about front-loading:

> "Rather than starting with the empirical results (as one would do in various indirect approaches), or with the training of subjects (as one would do on the neurophenomenological approach discussed above) this third approach would start with the experimental design. The idea is to front load phenomenological insights into the design of experiments, that is, to allow the insights developed in phenomenological analyses (modeled on Husserlian description, or the more empirically oriented phenomenological analyses found, for example, in Merleau-Ponty, or in previously completed neurophenomenological experiments) to inform the way experiments are set up." (2003, 91)

Taken in a mixed-method context, the idea is therefore to front-load concepts and distinctions from phenomenological analysis into the design of the qualitative and quantitative methods, and in this way theoretically frame the mixing of both methods. Throughout the article, we will give different examples of how to front-load phenomenology.

In Køster and Fernandez's (in press) own example, they front-load primarily Heideggerian, phenomenological concepts of "existentials" into their interview of people experiencing grief. These "existentials" refer to the essential structures of our being in the world, e.g., intentionality, selfhood, empathy, embodiment, temporality, spatiality, and affectivity. A quantitative example of front-loading is seen in cognitive science where the phenomenological distinction between "sense of agency" and "sense of ownership" is front-loaded into an experimental design using neuroimaging (Ruby and Decety, 2001; Chaminade and Decety, 2002; Farrer and Frith, 2002).

Being aware and taking responsibility for the phenomenological point of departure, i.e., commitments, concepts and distinctions, and front-loading it into the qualitative and quantitative methods, is therefore a necessary part of working with phenomenological mixed methods. Two of the most influential examples of how to work in such way comes from cognitive science and psychiatry.

Phenomenological Mixed Method in Experimental Settings

As mentioned in the introduction, neurophenomenology is one of the few examples of phenomenological mixed methods. At the core of how neurophenomenology is conducted is the qualitative interview method called the "explicitation interview" (Vermersch, 1994)—nowadays renamed to the "micro-phenomenological interview." What makes this interview method phenomenological is that Husserlian and other phenomenological ideas, distinctions (e.g., distinction between content and act of experience), and concepts (e.g., concepts of "pre-reflective experience" and "passive memory") are front-loaded into the concrete interview techniques of generating and analyzing data of a micro-experience.

Practically, this means that the researcher uses open how-questions (how would you describe your experience?) in the interview to help the participants: (i) reenact a past experience, (ii) suspend their beliefs and theories about this experience, (iii) redirect their attention from the content of the experience to the appearance of this content, and (iv) come into contact with the pre-reflective dimension and microdynamic processes of the experience—which are usually unrecognized, unnoticed, or concealed (see Petitmengin, 2006 for a detailed clarification). In the analysis, it means that the Husserlian and phenomenological ideas, concepts, and distinctions are applied in analyzing the descriptions of the participants' particular and lived experience. From such analysis, it is then possible to discover generic structures of experience (see Petitmengin et al., 2019 for a detailed clarification).

In a phenomenological mixed-method context, the aim is to correlate the data and analysis from the micro-phenomenological interview with the data and analysis generated from quantitative methods. This can be done in different ways depending on the specific micro-experience under investigation, the research questions, and the aim of the correlation procedure. These aspects would influence when the qualitative and quantitative data are generated and what quantitative methods are used. Some examples of quantitative methods include working with brain imagery to investigate the experience of illusory depth perception (Lutz, 2002; Lutz et al., 2002) and the experience of seizures for people with epilepsy (Le Van Quyen and Petitmengin, 2002; Petitmengin et al., 2006, 2007), working with experimental protocols such as the Rubber Hand Illusion (Valenzuela-Moguillansky, 2013) and decision tasks (Petitmengin et al., 2013), and working with physiological and cardiac measures to investigate the experience of surprise (Depraz, 2018).

In summary of these examples, the mixed method of neurophenomenology can vary according to the following parameters (Bitbol and Petitmengin, 2017):

a) Initiation: Starting with micro-phenomenological interview to identify experiential categories, or with quantitative measures to detect neuronal and physiological signatures.
b) Mode of identification: Identifying the experiential variables before and "front-load" them into the experimental and quantitative design or using micro-phenomenological interview to gather phenomenological descriptions after the experiment.
c) Level of temporal solution: Deciding at which time scale the correlation is looked for.
d) Level of genericity: Is the correlation sought at the generic (type) level between experiential structures and neural-physiological signatures or at the token level between singular experiences and their specific neural–physiological correlates?
e) Time analysis: If one wants to use new quantitative methods (e.g., intracranial Gamma-Band Mapping) to do real time analysis of the neuro-physiological signals and present the participants' and experimenters' with immediate (visual or auditory) feedback of the fine dynamics of this activity.

By reviewing two decades of literature on neurophenomenology, Berkovich-Ohana (2017) argues that neurophenomenology is appealing philosophically, but it is extremely difficult to implement experimentally, in both data generation and analysis. In some of the cases presented above, the method includes training the quantitative researchers and participants in Husserlian phenomenological methods such as epoché and phenomenological reduction. In neurophenomenology, there is therefore both the danger of "hyperphilosophizing" and that this mixed method primarily appeals to philosophers studying specific micro-experience. Micro-phenomenology has also been criticized for mis-representing phenomenology (Zahavi, 2011; Schmidt, 2018) and mis-construing a phenomenological understanding of the pre-reflective (Høffding and Martiny, 2016). That being said, it is nevertheless an increasingly popular and important method that begins to produce results and that we believe ought to influence and inspire phenomenological attempts at a mixed-method framework.

Phenomenological Mixed Method in Clinical Settings

In clinical work on schizophrenia, two phenomenological interview protocols have been developed to supplement standardized diagnostic systems such as ICD-10 and DSM-5. These interviews are called the Examination of Anomalous Self-experience ("EASE") (Parnas et al., 2005) and the Examination of Anomalous World Experience ("EAWE") (Sass et al., 2017). These protocols front-load a primarily Husserlian phenomenology into a semistructured qualitative interview design and a semiquantitative psychometric checklist to generate data of patients' subjective experience.

As an example of how phenomenology is front-loaded into EASE and EAWE, both interviews proceed from the phenomenologically secured insight of the existence of a minimal, pre-reflective self-awareness and work with the proposal that schizophrenia could be centrally grasped as a disturbance of this minimal self-awareness. The

phenomenological character of the EASE questionnaire can be directly inferred from many of its items, such as those in category 2, concerning "self-awareness and presence" (Parnas et al., 2005, 257). EASE also probes change in bodily experience (ibid), which can be seen as a continuation of the phenomenological insistence on the embodiment of subjectivity. EAWE further extends these insights, probing into changes in experience of the external world. This likewise flows out of a phenomenological orientation emphasizing the co-constitution of subjectivity and objectivity: certain changes in the experience of the world should reliably trace certain changes in subjectivity, found, in this case, in schizophrenia.

The motivation for working with such phenomenological mixed method is to obtain explanatory power in understanding, possibly diagnosing, and predicting schizophrenia, and thereby to expand the work on schizophrenia seen in the standardized diagnostic systems (Parnas and Henriksen, 2014), Henriksen et al. (under review). The interviews do so by prioritizing the patients' subjective experiences, which means that the focus is on the qualitative data of the patient's experience, which is generated while using the quantitative checklists and scoring sheets as a manual.

EASE and EAWE and their results have shown to be highly relevant both for diagnosis of schizophrenia and for the psychotherapeutic work involved in treating the condition. In relation to the latter, phenomenologically informed body-oriented psychotherapy (BPT) (Fuchs, 2005; Fuchs and De Jaegher, 2009; Fuchs and Schlimme, 2009; Koch and Fuchs, 2011; Fuchs and Koch, 2014; Fuchs et al., 2019) shows other ways to work with a phenomenological mixed method in schizophrenia research and clinical practice. In Galbusera et al. (2018), part of this work is conducted within an intervention framework where qualitative and quantitative methods are used to generate data sequentially before and after a BPT intervention. The methods for data generation include qualitative interviews, standardized symptom evaluation manuals, and Motion Energy Analysis software. They are used to understand and describe the therapeutic change processes for patients with schizophrenia using BPT, and the relation between the change processes and therapeutic results.

PROCESS: THREE CASES OF HOW TO USE PHENOMENOLOGICAL MIXED METHODS

Based on the inspiration, challenges, and few cases of phenomenological mixed methods presented above, we engaged in three different studies to investigate complex social phenomena using phenomenological mixed methods.

Understanding the Real-Life Experience of Joint Musical Absorption

Having established a phenomenological analysis of various forms of joint musical absorption (Høffding, 2019), we wanted to investigate the possible co-dependencies with heart-rate synchronization (HRS). Thus, concepts and distinctions from the phenomenological analysis were front-loaded into the design of the qualitative and quantitative methods.

The qualitative method used in the study was the approach of "phenomenological interview" where phenomenological analysis, commitments, ideas, and concepts are front-loaded into the interview design (Høffding and Martiny, 2016). The approach is a second-person, semistructured interview method with its own specific questioning and analysis techniques that use open "how" questions and specific strategies to co-generate detailed first-person descriptions of lived experiences. In this study, the interview focus was on how musicians experience their performance and how they experienced playing together.

Besides for HRS, there are many other biological and behavioral sources of synchronization we could have chosen to quantitatively investigate as co-determinants of joint musical absorption. Upham has conducted an activity analysis of music listeners' breathing synchronization and its coupling to music scores (Upham, 2018; Upham and Mcadams, 2018). Walton et al. have analyzed movement synchronization in music improvisation in a dynamical system framework (Walton et al., 2015, 2018); Swarbrick et al. have analyzed listeners' synchronized head bobbing (Swarbrick et al., 2019). Bishop et al. have analyzed both movement and eye-gaze synchronization in performers (Bishop et al., 2019), Bishop et al. (under review). In different settings, previous experiments have shown interesting correlations between personal relations and HRS in the context of shared experience in fire-walking ritual (Konvalinka et al., 2011) and choir-singing has also demonstrated strong couplings in heart rate (Müller and Lindenberger, 2011; Hemakom et al., 2016; Müller et al., 2018). We therefore chose to work with HRS.

The research strategy was consequently to cross-analyze the phenomenological interview data about the various experiences of joint musical absorption and the quantitative data measuring HRS (we used heart rate sensors produced by First Beat). To conduct the study, an interdisciplinary research team of biologists, psychologists, engineers, computer scientists, and phenomenologists collaborated with the musicians of "The Danish String Quartet."

In the study, we wanted to include concert performances where musicians were playing, while we investigate the relation between their experiences and HRS. As you cannot interrupt a string quartet with questions about their sense of absorption, while they are performing, we needed to develop a way to generate the data. The following process was developed: The musicians were playing with heart rate sensors on their chest underneath their shirts and so the quantitative data was generated during the musical performance. The phenomenological interview data was generated after the concert performance. This particular process raised a methodological question about how to mix the two data sets in a coherent and rigorous manner. This methodological challenge derives from the limitations of the phenomenological interview: even though it can disclose some experiential richness of past specific moments reflected on, it is not designed to hold a 1-to-1 relation with quantitative measures down to the millisecond[1].

[1] In this regard, the micro-phenomenological method, with its capacity to describe short periods of experience with great nuance, might be better suited to link the reflected experience with the quantitative data.

To face this methodological challenge, we recorded sound (Zoom H5 or H6) and video (Garmin 360 virb) during the musical performance. The sound recording was played back to the musicians individually allowing them to evaluate their experience of absorption during the performance in a self-rating application: they listen to a recording of their recently played concert while rating with one finger on a tablet using a sliding scale going from "distracted" to "very absorbed." This step ensures an automatic synchronization between the timing of the self-rating and the timing of the music. Moreover, since the recording of the heart rate is synchronized with the music in real time, we obtain a "bridge" to link the musicians' self-rated level of absorption and their heart rate.

After the musical performance, the musicians engaged in the self-rating sessions. The interviewing researcher could see the graphic representation of the results immediately after those sessions and used them to guide the qualitative interviews according to what was considered to be theoretically interesting. Already having in-depth knowledge of each of the DSQ musicians' phenomenology of absorption, relying on the ratings, Høffding could with a few questions ascertain how they experienced particular musical passages. In particular, he would enquire about each musicians' sense of the other ensemble members in selected moments, especially those of more intense forms of absorption. The data from the 360-degree video-recordings made it possible to contextualize the interpretation of the HRS, self-rating, and interviews. Among the relevant data provided by video-recordings were the musicians' behavioral and facial expressions.

Figure 1 shows an example of how the quantitative data was analyzed and visualized for one specific movement of a performance[2]. The similarities of beats per minute (BPM) across the musicians are clear in the graph, while the self-ratings of absorption are less homogeneous. Nevertheless, self-ratings converge around second 450, where a clear drop of absorption can be observed for two of the musicians (Rune and Frederik S.), and to a lesser degree for a third musician (Frederik Ø.). Just before this drop, there is also an absorption peak for all four musicians.

The qualitative data was analyzed in a descriptive manner. This meant that the recorded interviews were transcribed and from the transcription the data was coded and structured in order to identify experiential categories of joint musical absorption.

In the interpretation of the different analyses, the qualitative data analysis was contrasted with the quantitative data analysis for enriching the self-rating data. Here we see that two of the musicians describe that the drop in the absorption rating represents a transition from the fourth to the fifth variation in the third movement of Beethoven's 5th string quartet (opus 18), which is a compositional change from something intimate, slow, and *piano* to something much quicker, *forte*, merrier, and almost silly.

[2]The raw data (audio recording, self-rating app scores, and the First Beat sensor data) is available at: https://www.uio.no/ritmo/english/news-and-events/events/musiclab/2021/dsq/.

In quantitative terms, this development in the composition manifests itself in an increased amount of bodily motion, which can account for the BPM development in **Figure 1**. From the video recording, we see a break of around 0.5 s preceding the fifth variation, which enters in a sudden and energetic way after a clearly audible joint, deep, and fast in-breath. Two of the musicians reported the experience of "merryness" as associated with some level of ironic self-distance, which is absent in the intimate, *piano* variation preceding it. The violist, Asbjørn Nørgaard, agrees on the funny and almost silly character of the fifth variation. However, he does not link it to a drop in absorption since he considers himself equally absorbed in the funniness. These experiential reports help interpreting the different self-ratings for the three musicians. The fourth musician did not report a significant experience about the fifth variation.

The slightly contorted and even smiling faces of two of the musicians during those crucial seconds around the fifth variation, observable in the video recordings, contribute to contextualize the mixed analysis of the datasets. One might conclude that the change in the composition from the fourth to the fifth variation induces a drop in experienced absorption, which, perhaps not surprisingly, seems to generalize to the conclusion that musical genre and quality strongly impact the sense of joint absorption of the performers. Further, one can conclude that the self-ratings on their own can be deceptive and are best understood when constrained by interviews concerning those ratings.

Finally, as an exploratory pilot experiment, the current study also shows the complexity of integrating qualitative (interviews) and physiological (HRS) aspects of musical absorption. The fire-walking study (Konvalinka et al., 2011) that inspired the study on joint musical absorption showed a clear correlation between HRS and personal relations between actors and spectators in the ceremony. Part of its success was the relative simplicity of the task of crossing the burning coals coupled with the audience being stationary, all directly impacting the BPM and hence the HRS. When performing, all four musicians, however, are almost constantly moving at different paces: we need more research to identify ways to separate analyses of interview, self-rating, music score, quantity of motion, and HR before a meta-analysis and interpretation can more conclusively answer if heart rate synchronization plays a role in joint musical absorption.

How to Create and Investigate Positive and Negative Experiences of Joint Actions in CP

In the study on joint actions involving people with CP, we wanted to answer the research question: how do people with CP, who have problems with bodily coordination, experience positive and negative joint actions? (Toro, 2020; Toro and Martiny, 2020), Martiny et al. (under review).

To answer this question, we needed to design a study where we could create positive and negative joint actions, investigate the data of bodily coordination and positive or negative experiences of joint interactions involving persons with CP, and then compare the data with similar data from interactions involving persons without CP (a control group).

FIGURE 1 | Graphs of post-concert self-ratings for absorption during a part of a particular concert in Denmark (left panel) and the associated HR profiles for each quartet member recorded during the actual concert. The y-axes represent (arbitrary) values on the rating scale, and heart beats per minute (BPM), respectively. The x-axes represent time in seconds. Note that the y-axes for each plot are scaled to the individual rating and HR curves for purposes of visual presentation.

From previous phenomenologically informed qualitative work with CP (Martiny, 2015a,b), we knew that CP interactions are experientially different depending on who the people with CP are interacting with. So, we wanted to include CP interactions where one person with CP would perform joint actions with one person from three different groups of non-CP participants: (1) relatives, (2) therapists (stranger group #1), and (3) random strangers (stranger group #2). In the control setting, we included two persons without CP to perform the same joint actions.

We also knew from the previous work that everyday joint actions such as shaking hands, giving/receiving an object, or carrying an object together are very challenging for people with CP. So, depending on who the person with CP would interact with, the experience of such joint actions would be either positive or negative. In the experiment, we therefore developed six daily, hand-to-hand joint action exercises, which the two participants would perform sitting down at a table in front of each other.

In phenomenology, our mode of being-in-the-world is described as structured by our embodiment and primordially action-oriented. This means that corresponding with our worldly interests and our bodily skills and competences, we perceive objects as invitations for specific actions and we coordinate our bodily movement according to these perceptions (Merleau-Ponty, 2012). However, we also perceive other people as affording specific bodily responses and ways of engaging. This phenomenological analysis resonates well with Gibson's theory of affordances (Gibson, 1979).

Based on this analysis and theory, we therefore decided to track the participants' eye movements, as well as their bodily movements during the joint-action exercise. This quantitative data was generated using eye-tracking glasses (Tobii Pro Glasses 2) that the participants wore, so that both their eye focus and areas of interest were recorded. In addition, a Kinect (v1) camera was set up in the room to record 3D video data of the participants' bodily motion and movements.

The eye-tracking devices would also allow us to determine what regions of the environment seemed more attractive or relevant for the interactors during the interactions. We would then be able to analyze the predominant eye focus of the participants during the interaction—thus providing a signal of the person's attitude toward the interaction. For the idea of including the participants' attitudes in the analysis, we front-loaded Husserl's notions of personalistic and naturalistic attitudes (Husserl, 1989; Toro and Martiny, 2020).

Immediately after the joint action exercises were performed, we conducted a 15–20-min interview together with both participants to generate qualitative data of how the participants experienced the interactions. The interview was conducted using the approach of a "phenomenological interview," as described above. In this study, the interview focus was on how participants with and without CP experienced the situation of acting together in the joint action exercises, how they experienced their own actions and the actions of the other participant, and how they experienced the interaction when functional challenges occurred.

To be able to conduct such a phenomenological mixed-method study, with both qualitative and quantitative data generation, we were an interdisciplinary team of philosophers, psychologists, and computer and cognitive scientists. For the analysis of the quantitative eye-tracking and bodily motion data, we imported the data and analyzed it using MATLAB. The

quantitative analysis employed classical statistical models such as linear mixed-effects, full, and null models.

The quantitative analysis showed that during the interaction, the strangers looked significantly less at the face of the person with CP compared with relatives and therapists. The control group looked at each other's face much more than participants in the CP interactions and was also significantly faster in performing the exercises. Despite functional challenges, all CP interactions completed the exercises successfully and it took them approximately the same time. In the most demanding exercises, CP–stranger interactions were as quick or quicker than CP–relatives and CP–therapists. Also, in the CP–stranger and control groups the participants responded quicker to the facilitator's instructions and moved quicker toward the point of interaction, than CP–relatives and CP–therapists.

The qualitative data was analyzed in a descriptive manner. This meant that the recorded interviews were transcribed and from the transcription the data was coded and structured in order to identify experiential categories. The qualitative reports were categorized according to the participants' experience of the interaction in general (as positive or negative), their experience of the task, of the other person, of the situation, and of themselves. In the analysis, we found elements that allowed us to identify common categories of positive as well as negative experiences. Positive experiences were described as "natural," "open," "attuned," "habitual," and "calm." Negative experiences were described with terms like "unnatural," "alert," "transgressive," "functional," "hesitant," and "correct." Overall, CP–stranger interactions were experienced negatively, while CP–relatives, CP–therapists, and controls had positive experiences of joint actions.

In the two separate data analyses, we observed that most cases of CP–stranger interactions showed relatively high levels of coordination and goal accomplishment but were experienced as negative. We also observed that even though the CP–therapist and CP–relative interactions were functionally more challenging, they were experienced much more positively. To make sense of this complexity, we triangulated the two datasets with the front-loaded and a phenomenologically enriched theory of affordances to disclose the fundamental structures of the complexity of joint actions (see **Figure 2**).

The theoretical framework allowed us to mix and interpret the data into a unified account of the phenomenon of joint actions involving interactions with people with CP. Our findings contest current embodied and enactive accounts of joint actions, according to which there is a direct relation between the functional and experiential dimensions in joint actions. Our study showed the different ways that functionality and affectivity are interwoven in joint actions and how they are mediated by a third dimension of joint action. We propose that this third dimension is the openness of the system constituted by the participants in a joint action [see Martiny et al. (under review)].

Our proposal is phenomenologically inspired, as it reflects the complexity of social interactions—not only as relations between living—physical—bodies, objectively describable in terms of bodily coordination—but primordially, as a relation between lived bodies, embedded in socially and culturally rich contexts. Indeed, our experience of our own bodies, of the other person, of the situation, and of the interaction depends on more than just bodily coordination and performing specific tasks successfully.

How to Change and Investigate Prejudicial Attitudes Toward Physical Disability

In the study of reducing prejudice toward people with physical disability, we wanted to test our hypothesis that embodied engagement as contact reduces prejudice toward physical disability by changing attitudes to be more positive toward people with physical disabilities. To test this hypothesis, we needed a design where we could investigate peoples' explicit (conscious) and implicit (unconscious) processes of attitude formation before and after an embodied engagement-based intervention.

As it turns out, contemporary ways of thinking about theater in theater and performance study is very much in line with phenomenologically informed social cognitive science. Theater is described by three notions, namely, (1) embodiment, (2) engagement, and (3) transformation. The notion of "embodiment" is used as a way of understanding the "affective" impacts that the performance has on the audience (Thompson, 2009; Nicholson, 2011), where the audience experiences "being kinaesthetically moved" (Fenemore, 2003). The notion of "engagement" is used to emphasize that in theater performance, performers and audience are in a shared, participatory, and immersed dialogue (Shepherd, 2006; Shaughnessy, 2012). The idea and notion of "transformation" describe how the engaged performance creates "effect" of social change through its embodied "affect" (Nicholson, 2005; Thompson, 2009). Thus, it seemed to us that theater would allow for embodied engagement and that it had the potential to create change.

A large team of both researchers (philosophers, psychologists, and researchers in performance studies and cognitive science) and practitioners (theater director, actors, dramaturgs, scenographer, and musician) collaborated to develop a specific embodied engagement intervention using a theater performance. The theater performance was developed as an autobiographical stage performance about a 28-year-old man, JN, who lives with quadriplegic cerebral palsy (CP) and has a speech impediment, thus displaying group salience both visibly and audibly at first encounter.

To measure the engagement-based intervention's effect on the audience, we wanted to investigate their attitudes toward physical disability before, during, and after the intervention. To do so, we designed a mixed-method research process consisting of both explicit (self-reporting) quantitative and qualitative methods, as well as implicit (behavioral) quantitative measures. The different methods were testing for both the successful achievement of embodied engagement and attitude change as part of the prejudice reduction.

The quantitative data was generated before and after using a 7-point Likert scale quantitative questionnaire and an Implicit Association Test (IAT), and it was generated at a restricted research test area close to the theater stage. The IAT test is a standardized test used to measure implicit attitudes, which means that we could not front-load phenomenological

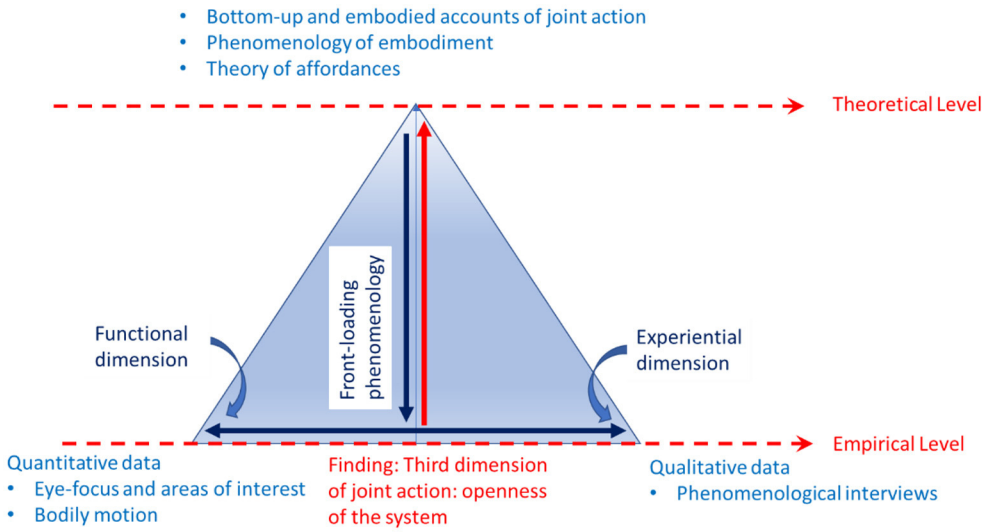


FIGURE 2 | Diagram of the triangulation in the joint action experiment involving people with CP. The triangle depicts vertically the relations between the theoretical and empirical levels, and horizontally the relation between quantitative and qualitative data. At the theoretical level, the experiment was front-loaded by three main sources: (1) bottom-up and embodied accounts of joint action, (2) insights from the phenomenology of embodiment, and (3) the theory of affordances. At the empirical level, the relation between the functional and experiential dimensions is studied based on quantitative and qualitative data (horizontal blue arrow). From this analysis, an interpretation at the empirical level emerges, which suggests a different account of joint actions than the one offered by current theory (red vertical arrow).

analyses, concepts, and distinctions into the test. However, the quantitative questionnaire was developed by choosing questions from previous quantitative questionnaires on physical disability and changing the way the questions was asked in order to correspond to the "how" questioning techniques coming from the phenomenological interview.

During the performance, an interactive questionnaire was also developed to generate quantitative data of the audience's behavior in forming their attitude toward physical disability. We wanted to see if their engagement toward JN on stage continued throughout the performance. So, we front-loaded the definition of engagement from Satne and Roepstorff (2015) into the design of an interactive questionnaire about JN, which was conducted within the actual performance. Satne and Ropestorff define engagement as an affective, emotional, and reciprocal we-experience in which one is committed to the other as a person. If the audience were personally committed to JN, we expected that they would then continue to answer the questions we ask them thought the performance. We therefore measured how long it took the audience to answer the questions of the interactive questionnaire, the number of audience members who answered, and whether they made answer revisions. The questions were designed based on concrete scenes that the audience had just experienced, and the answers were submitted on a 10-point Likert scale by using a mobile answering device that was placed in the participants' seats. The answers were shown in "real time," anonymously on a projection wall on stage that everyone could follow.

Data from most of the audience ($n = 2604$) was generated using the interactive questionnaire, but only parts of the audience were recruited before (pre-group) and after (post-group) the performance and given the quantitative questionnaire and the IAT test. For comparison, a control group ($n = 505$) that did not see the performance was given the quantitative questionnaire. A focus group ($n = 30$) was also recruited and given the quantitative questionnaire and the IAT test after the performance. Fifteen participants of the focus group were chosen for follow-up qualitative interviews. The interview was conducted using the approach of "phenomenological interview," as described above. In this study, the interview focused on how the audience experienced both JN on the stage and the theater performance in general and their experience of answering the interactive questionnaire.

The quantitative and qualitative data were analyzed separately. The quantitative analysis concerned statistical comparison of the data before, during, and after the performance. Here we employed classical statistical models such as a series Kruskal–Wallis H-test for the quantitative questionnaire, an independent-sample Mann–Whitney *U*-test conducted in SPSS for the IAT test, and the interactive questionnaire was analyzed descriptively.

The qualitative data was analyzed in a descriptive manner. This meant that the recorded interviews were transcribed and from the transcription the data was coded and structured in order to identify experiential categories. The qualitative reports were categorized according to three overall categories of experience: (1) the experience of the intervention, (2) experiences of a nuancing and reflective effect, and (3) experience of the attitude formation.

The qualitative data showed that the performance changed the audience's attitudes toward people with physical disability from being objectifying and prejudicial to being humanizing, personalistic, and inclusive. The interactive questionnaire

showed that the audience was highly engaged during the performance, with a small amount of people in the audience that changed their answers. The IAT test showed no change in attitudes after the performance, whereas the quantitative questionnaire showed both significant positive changes in the attitudes of people in the audience as well as the effects of a decrease in positive attitudes.

The decrease in positive attitudes was surprising to us, since our hypothesis was that embodied engagement as contact would reduce prejudice toward physical disability by changing attitudes to be more positive toward people with physical disabilities. To understand this apparent conflict between our hypothesis and data, we used a triangulation strategy to develop a meta-interpretation of the different analyses.

The front-loaded and phenomenologically based, second-person theory of engagement was used to combine and mix the quantitative and qualitative datasets into one account of prejudice reduction. In this account, reduction is not seen as a matter of decreasing negative attitudes or increasing positive attitudes. In contrast to Contact theories, prejudice reduction in our account is understood as the nuancing of attitudes. In the performance, this nuance effect is initiated by a self-reflection process where people in the audience become aware of the act of forming their own attitudes. This attitude formation is highly influenced and intensified by the social setting of the theater.

Given this interpretation, we reformulated our initial hypothesis of focusing on positive attitude toward physical disability and presented this reformulation in terms of the Engage, Nuance, and Attitude formation (Enact) Hypothesis. The Enact hypothesis states that to change prejudicial attitudes, interventions should be designed so that persons involved become highly engaged with the attitude object (e.g., personalized outgroup member), engaging on both an embodied, affective, behavioral, and social level. Further, we suggest that the goal of prejudice reductions should not be thought of as changing either positive or negative attitudes, but as a nuancing of attitudes.

According to the Enact hypothesis, the reduction of prejudice occurs due to the increased embodied engagement with people with physical disability, which creates an explicit and conscious nuancing and self-reflective effect. This also explains the lack of significance in the IAT test, since such tests target "automatic" and "implicit" associations that operate at a lower (un)conscious level of attitude formation and change.

GUIDANCE: STEPS, DECISIONS, AND STANDARDS

What does the inspiration from the previous work combined with the lessons learned in our studies mean for a researcher who wants to engage in phenomenological mixed methods? In this section, the aim is to collect the insights, provide guidance, and help those interested in phenomenological mixed methods to make consistent decisions through the different methodological steps.

As we have seen in all the different cases and examples, when working with phenomenological mixed methods, the first part is to clarify one's phenomenological frame and point of departure, i.e., commitments, theories, analyses, concepts, and distinctions. Here the steps to take and decisions to make regard how one will front-load phenomenology into one's research question and mixed-method design. The second part, as we have seen, concerns what this phenomenological frame and front-loading means for how we generate the qualitative and quantitative data (tier one). The third and last part concerns how the front-loaded phenomenology informs the analysis and interpretation of the qualitative and quantitative data (tier two).

In the following section, we will go through the three different parts, clarify the steps and decisions within each part, and end by discussing what defines consistency in the steps and decisions.

The Framing: A Phenomenological Point of Departure

As argued above, phenomenology provides a good starting point to mix qualitative and quantitative methods. Here we can avoid the challenge of "hyperphilosophy" by front-loading phenomenology into the mixed-method design, instead of getting caught up in methodological orthodoxy.

The first step is therefore to use phenomenology to guide one's research question. This means that one should mix into the question both subjective (first-person) and objective (third-person) aspects of the experience one is trying to investigate. That could be questions including pre-reflective and microdynamic processes of experiences and their neural signatures, joint experiences (e.g., joint musical absorption and joint actions) and their corresponding bodily and physiological aspects (heart rate, eye gaze, and bodily motion), or experiences of prejudice and their implicit and behavioral components.

As we see in the different cases and examples, the specific motivations for why we want to investigate a specific experience differ. The next step is to clarify one's motivation, since the design of the phenomenological mixed method will vary according to the motivation and so will the interpretation of the data. Using Venkatesh et al. (2013) as inspiration, we can group the different motivations for working with phenomenological mixed methods into three categories:

1. *Strengthening:* An approach designed for using both qualitative and quantitative methods to strengthen the understanding of a specific experience. This is seen in the case of neurophenomenology, which uses qualitative and quantitative methods with two motivational reasons in mind:

a. Corroboration: To verify the findings from one type of data with data from the other type.
b. Compensation: To utilize one method and its data to compensate for the weaknesses of the other type of method.

2. *Improvement:* An approach designed for using both qualitative and quantitative methods to develop a richer understanding of an experience. This we can do either through:

a. Expansion: By utilizing the one type of study to expand the understanding of the findings of the other type. This is the motivation behind our study of musical absorption, where the aim is to expand already established qualitative analysis of joint musical absorption by measuring quantitative heart-rate data and cross-analyze both data sets.
b. Developmental: By using the one type of study to develop research questions, hypotheses, and understanding for the other type. This is seen in the case of EASE and EAWE where qualitative method and data (i.e., the patients' subjective experience) are prioritized in order to develop better understanding of schizophrenia within the quantitative, standardized diagnostic system.

3. *Holistic:* To use both qualitative and quantitative methods to develop a more holistic understanding of an experience, either through:

a. Complementarity/divergence: By using the mixed methods to gain complementary or divergent views on the same experience. This is the motivation behind our study of prejudice against physical disability, since we used both qualitative and quantitative methods to gain different views on prejudicial experience.
b. Completeness: To use the mixed methods to provide a complete picture of an experience. This is the motivation behind our study of joint actions involving people with CP, where we used both quantitative and qualitative methods to get a more complete picture of both the functional and affective aspects of joint actions.

After one has decided on either strengthening, improving, or providing a holistic understanding of a specific experience, the third and last step in framing one's research is to choose which type of phenomenological design to work with. In order to take this step, one must answer the following questions: What will be the significance of the strengthened, improved, or holistic understanding of the experience and why should one engage in phenomenological mixed methods for providing such understanding?

If the significance and scope relate primarily to the theoretical and experimental research field within which the research is done (e.g., research in joint actions), one can choose to work with a basic design. This can be a basic study and/or experiment, as we saw in our study of joint actions in CP, and which we saw in many of the mixed-method cases in neurophenomenology.

However, the design combination of phenomenology with qualitative and quantitative methods opens up possibilities for case studies and researching experiences in real-life contexts (e.g., musicians playing), for working with interventions (e.g., BPT intervention) and changing experiences (e.g., prejudicial attitudes) and for including the lived experience of patients to transform diagnosis and therapy (e.g., EASE and EAWE). Many of the productive examples of phenomenological mixed methods in health and clinical settings work with a transformative aspect, since they aim to improve or provide better therapy and healthcare (see Zahavi and Martiny, 2019; Toro and Martiny, 2020). If the significance and scope of one's research relates to case studies and interventions, including participants' lived experience (participatory research) or transformative matters, it makes it a type of advanced phenomenological design (Fetters et al., 2013).

After being clear on one's phenomenological point of departure, research question, motivation, and type of design, one has to figure out how to phenomenologically inform the data generation.

Tier One: Phenomenologically Informed Data Generation

The first step in tier one is to clarify how phenomenology will be front-loaded into the qualitative and quantitative data generation. As we have seen, when it comes to the qualitative data generation this means working—in all the different cases and examples—with some version of the "phenomenological interview." Depending on the experiences one is investigating, one can front-load different phenomenological analyses, concepts, and distinctions into the interview and generate the qualitative data in different ways. This means that there will be different possibilities for working with the interview that apply different framing, interview foci, questions, and techniques.

That being said, there is currently a methodological gap in how phenomenology is and can be front-loaded into other qualitative methods in a mixed-method context. What does it mean for phenomenology to be front-loaded into and inform for example: participant observations, video analysis, archival investigations, or discourse analysis? In Martiny et al. (2016), we discuss and give examples of how one can work phenomenologically with multimedia within qualitative methods and in our study of joint musical absorption we also used video analysis. However, to develop clear guidelines for how phenomenology could be front-loaded into a variety of qualitative methods would be a fruitful way to improve research within phenomenological mixed methods.

When it comes to the quantitative data generation, the decision is not about choosing one particular approach of how to front-load phenomenology into one quantitative method. Rather, as we have seen, there are many different examples of quantitative methods that phenomenology can be front-loaded into. For example, phenomenology can be front-loaded into the application of brain imagery, psychological tests and experimental protocols, quantitative questionnaires, standardized manuals, eye-trackers, bodily motion measures, motion energy measures, heart rate, and other physiological and cardiac measures. It is easier to make the decision about which quantitative methods to use by keeping the focus on the mixed aspect of the research. What is unique about phenomenological mixed methods is not which quantitative method one uses but how the qualitative and quantitative data generation processes are mixed and integrated into one's phenomenological research inquiry.

With inspiration from mixed-method research (Creswell et al., 2003), the next step to take in designing a mixed data generation process requires answering questions like: will the qualitative and quantitative data be generate more or less at the

same time (concurrent) or in different phases, over a period of time (sequential)? Will the qualitative or quantitative data generation be prioritized as equally important, or will one be prioritized over the other, due to, e.g., practical constraints of data generation or the need to understand the one type of data before moving on to the other? How and when will the qualitative and quantitative data be integrated? By analyzing one data set before moving on to the next (connecting); by merging the two data sets in the interpretation; or by embedding one data generation process into the generation of the other datasets?

Depending on how one answers these questions, the data generation processes can be designed differently. As we see in neurophenomenology, one can start with the micro-phenomenological interview to help inform and/or interpret the quantitative data generation (e.g., brain imagery). One can also generate the quantitative data first and then conduct the micro-phenomenological interview afterward. In this way, the quantitative data will help inform and/or interpret the qualitative data. The first approach is an exploratory design, and the second is an explanatory design, but similar for both is having one research track where the data is generated sequentially and then connected to one another using the phenomenological frame.

In our studies of joint actions in CP and reduction of prejudice toward physical disability, we see that it is also possible to generate the qualitative and quantitative data more or less concurrently, i.e., during a similar timeframe. Here the phenomenological interview is used to generate the qualitative data and different methods (eye-tracking, bodily motion measures, IAT test, and questionnaires) are used to generate the quantitative data. Characteristic for such approach is that the two data generation processes are not dependent on one another. This means that one will have two tracks where the qualitative and quantitative data are generated in parallel. Using the phenomenological frame, the data can then be triangulated and merged in the interpretation of the data (tier two).

In EASE and EAWE, as well as in our study of joint musical absorption, we see that the qualitative and quantitative data generation is also done more or less concurrently. However, in the former, the qualitative data of the patients' experience are prioritized and the quantitative checklist and scoring sheets are applied as a manual for the qualitative interview. In the latter study, the phenomenological interview is embedded within a largely quantitative generation process. In this way of working with the mixed method, there will be one track of data generation where one of the qualitative and quantitative data processes crosses over and is embedded within the one. The phenomenological frame ensures that this crossover is theoretically coherent and therefore can occur.

Using inspiration from mixed-method research (Creswell, 2003; Creswell and Plano Clark, 2007), we can categorize the following four major types of data generation processes, where the phenomenological frame integrates qualitative and quantitative data: Explanatory, exploratory, triangulation, and embedding (see **Figure 3**).

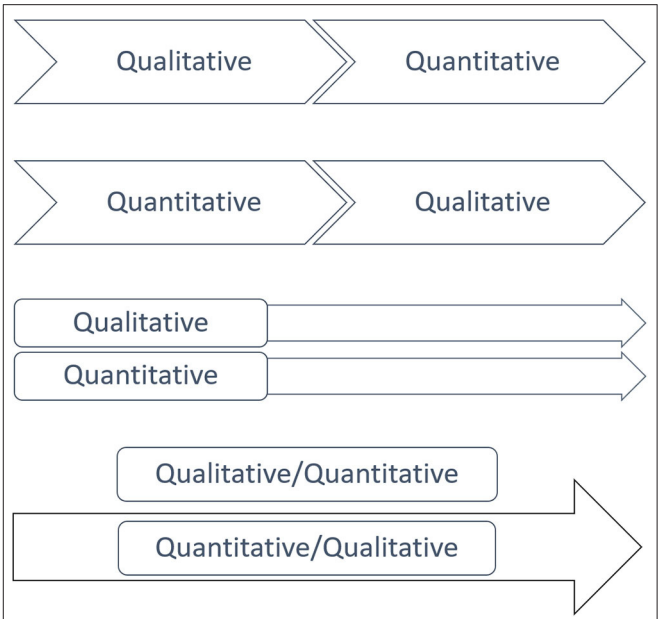


FIGURE 3 | A description and illustration of the four types of data generation processes in phenomenological mixed method.

Tier Two: Phenomenologically Informed Analysis and Interpretation

After having designed the data generation process, one must figure out how to analyze and interpret the data phenomenologically. Here, the first step follows from one's decisions in tier one. If one decided to generate the data sequentially (explanatory or exploratory) or in parallel (triangulation), the analysis of the qualitative and quantitative data will be done separately following either the sequential or parallel processes. This way of analyzing the data is seen in neurophenomenology and in our two studies on physical disability.

If one instead decides on embedding the data generation, the analysis will be one of conversion, where one form of data is converted into another form. This conversion can occur through quantification, as seen in EASE and EAWE, where the qualitative data (patients' experiences) are converted into quantitative data (numbers on a scoring sheet). It can also occur through qualification, as seen in our study on joint musical absorption, where quantitative data (self-rating) is converted into qualitative data (descriptions and narratives of the related experience). In our study of joint musical absorption, however, we also did separate qualitative and quantitative analyses, which means that this is a case of multimixed analysis, where one can use a combination of both separate and converted analyses in the process.

For the qualitative part of the data analysis—whether it is done separately or through conversion—phenomenology plays an important and explicit role. As seen in all the different cases and examples, the phenomenological analyses, concepts, and distinctions are applied in working with the recorded

and transcribed interview data. The analysis process concerns coding and structuring the data in order to identify experiential categories and/or themes. For the quantitative part of the data analysis, we see different ways for phenomenology to inform the analysis in the cases and examples. The front-loaded phenomenological analyses, concepts, and distinctions help in deciding, for example, which quantitative data sample to focus on and which statistical models to use in analyzing this sample.

After having conducted the analysis, the next step is to interpret and make sense of the separated, converted, or multimixed analysis. one has for working with a phenomenological mixed method will influence this interpretation, since one will be looking for a strengthened, improved, or holistic understanding of a specific experience. This means that a fundamental part of the sense-making process in the interpretation includes keeping one's phenomenological frame, point of departure, and research questions in the foreground.

In addition to this suggestion, Teddlie and Tashakkori (2009, p. 289-293) provide some more general guidelines for mixed method sense-making. This includes separating one's research question into sub-questions, so that the relevant results for each sub-question can be summarized and examined. The questions can, for instance, be separated into the qualitative and quantitative research tracks (data and analysis). This exercise will provide some tentative interpretations and answers to the questions, which should then be mixed, i.e., compared, contrasted, combined, or the difference between them should be explained. The overall aim of this mixing is for the different interpretations and answers to the sub-questions to be integrated into one meta-interpretation.

For a phenomenological mixed method, the meta-interpretation should provide a generalized understanding of the structures of the experience in question (see commitments in section Phenomenological Foundations and Commitments). This means that although some partial interpretations and answers within the sense-making process might refer to particular experiences, the meta-interpretation should end up with generalized descriptions and understanding. As seen in the different cases and examples above, one is therefore able to both understand the experiences of particular patients, persons with disability and musicians, and what this means for understanding experiential structures in relation to, e.g., schizophrenia, joint actions, reduction of prejudice, and joint musical absorption. In a phenomenological mixed method, the interpretations and answers will therefore include a continuum of both particularity and generalization.

The meta-interpretation and the provided research answers of the phenomenological mixed method should be assessed in terms of the validity and quality of the sense-making processes. According to Tashakkori and Teddlie (2008) and Teddlie and Tashakkori (2009), this refers to the quality of the design and procedures, and the rigor and transferability (i.e., replication) of the interpretation and research results. In relation to a phenomenological mixed method, we propose that such validity and quality can be understood in terms of the "performative consistency" (Petitmengin and Bitbol, 2009) and the "phenomenological consistency" (Høffding and Martiny, 2016) of the sense-making process. This means that good phenomenological mixed-method research is conducted when there is a high degree of the following forms of consistency (see also **Figure 4**):

1. *Performative consistency*: The degree to which there is consistency between the three parts of the phenomenological mixed method, i.e., the phenomenological frame (commitments, theories, and concepts), tier one (the qualitative and quantitative data generation), and tier two (the mixed analysis and meta-interpretation).
2. *Internal phenomenological consistency*: The degree to which it is possible in the meta-interpretation to provide clear and coherent descriptions and explanations of the different qualitative and quantitative data and their relation.
3. *External phenomenological consistency*: The degree to which it is possible for the meta-interpretation to work with and against already established theories and understanding of the specific experiences in question. External phenomenological consistency is related to the methodological steps of "intersubjective validation" (Varela and Shear, 1999, 10) and "intersubjective corroboration" (Gallagher and Zahavi, 2008, 29–31).

In relation to neurophenomenology, Petitmengin and Bitbol describe "performative consistency" as a "validity in action," which refers to the reproducibility of the method. They aim at prescribing a manual-like procedure for how to "correctly" conduct neurophenomenology. In Høffding and Martiny (2016), we acknowledge the overall idea of consistency between one's theories, methods, data, analyses, and interpretation but criticize the manual-like interpretation of performative consistency. In relation to phenomenological mixed methods, as we have seen in the different cases and examples, there is no one "correct" way to conduct the research, but many ways to mix methods phenomenologically. However, there should be a high degree of relational consistency between the three parts, i.e., one's phenomenological frame, how phenomenology is front-loaded and informs the methods used for data generation, and how the data is analyzed and interpreted based on the phenomenological frame and mixed methods.

This performative consistency will ensure that one's research can acquire a high degree of internal consistency. As we see in our three studies, for example, when complexity arises in comparing the qualitative and quantitative data or it seems that there are apparent conflicts between hypothesis and data, the phenomenological frame and theories can help in making clear and coherent sense of the data. In the two studies on physical disability, we also tried to create a high degree of external phenomenological consistency by not limiting the interpretations to already established theories of joint action and prejudice reduction. We also applied our joint action account in a rehabilitation research context to develop better-personalized healthcare for people with physical disabilities (Toro and Martiny, 2020) and develop a new theater performance for reducing prejudice for people with depression.

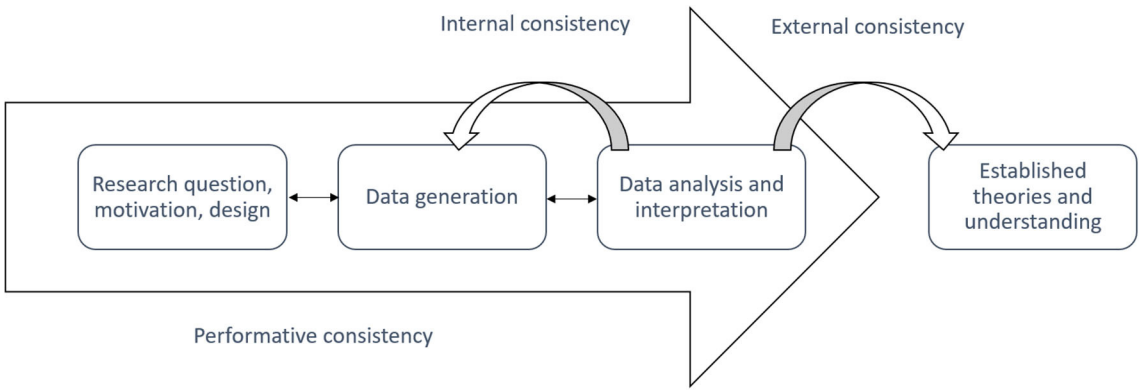


FIGURE 4 | An illustration of the three forms of consistency in phenomenological mixed methods. Performative consistency refers to the consistency between the three parts of the phenomenological mixed method: Phenomenological frame, tier one, and tier two. Internal phenomenological consistency refers to how the interpretation relates to the data, and external phenomenological consistency refers to how the interpretation relates to already established theories and understanding of the investigated experience.

CONCLUSION

In this article, we aimed at unearthing and making explicit important methodological considerations underlying a phenomenological mixed method, to guide researchers through the difficulties of studying experience using both qualitative and quantitative methods.

By framing the mixed-method research phenomenologically and beginning with a phenomenological point of departure, the research will proceed according to clearly established commitments that avoid the "whatever works" rationale and help the researcher to guarantee the consistency of their research.

As we propose, in applying the phenomenological frame one avoids "hyperphilosophizing" by front-loading phenomenology into the mixed method design, rather than getting caught up in methodological orthodoxy. This means that one should figure out what aspects of phenomenological analyses and theories are front-loaded, how they are front-loaded, and what this means for one's research question, motivation, type of design, data generation, analysis, and interpretation.

We have endeavored to show that there are different ways in which qualitative and quantitative methods can be mixed phenomenologically. We have developed the three-fold structure (the phenomenological frame, tier one, and tier two) for conducting phenomenological mixed-method research as a guideline through the landscape of possibilities available. The aim of the three-part structure is for those interested in phenomenological mixed methods to take steps and make decisions that are performatively and phenomenologically consistent.

DATA AVAILABILITY STATEMENT

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

ETHICS STATEMENT

Høffding's study with the DSQ was reviewed and approved by the Norwegin Center for Research Data under number 613262. For the other studies, ethical review and approval was not required in accordance with the local legislation and institutional requirements. 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

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

FUNDING

KM was supported by Enactlab S/I, JT research was supported by Colciencias, grant 756–2016, and SH research was supported by Norges Forskningsråd, grant 262762.

ACKNOWLEDGMENTS

Conducting phenomenological mixed-method research is a collaborative effort, so in relation to the three examples of phenomenological mixed method, we want to thank Kasper Løvborg Jensen, Jonna Katariina Vuoskoski, Eigil Yuichi Hyldgaard Lippert, Andreas Roepstorff, and Sebastian Wallot for their invaluable work on the musical absorption study. We want to thank Mikkel D. Justiniano, John Paulin Hansen, and Per Bækgaard for their important contribution in the technical aspects of the joint action experiment. Also, thanks are given to the people at the Enactlab, Jacob Nossell, David Eskelund Nielsen, and Asger Juhl for their help in developing and performing the

joint action experiment. Lastly, we want to thank Helene Scott-Fordsmand, Andreas R. Jensen, Asger Juhl, David E. Nielsen, and Thomas Corneliussen and his theater team for their incredible work in developing and conducting the study on prejudice toward people with physical disability. In relation to this, thanks should also be given to the graduate students who helped in gathering data, namely, Sofie B. Jørgensen, Kenni Antonsen, Kristoffer F. Dalsgård, and Line M. Møller.

REFERENCES

Berkovich-Ohana, A. (2017). *Radical Neurophenomenology: We Cannot Solve the Problems Using the Same Kind of Thinking We Used When We Created Them*. Constructivist Foundations, 12, 156–159.

Binder, J., Zagefka, H., Brown, R., Funke, F., Kessler, T., Mummendey, A., et al. (2009). Does contact reduce prejudice or does prejudice reduce contact? A longitudinal test of the contact hypothesis among majority and minority groups in three European Countries. *J. Personality Soc. Psychol.* 96, 843–856. doi: 10.1037/a0013470

Bishop, L., Cancino-Chacon, C. E., and Goebl, W. (2019). Moving to communicate, moving to interact: patterns of body motion in musical duo performance. *Music Percept.* 37, 1–25. doi: 10.1525/mp.2019.37.1.1

Bitbol, M., and Petitmengin, C. (2017). "Neurophenomenology and the microphenomenological interview," in *The Blackwell Companion to Consciousness*, eds S. Schneider and M. Velmans (Wiley Blackwell, Oxford), 726–739. doi: 10.1002/9781119132363.ch51

Chaminade, T., and Decety, J. (2002). Leader or follower? Involvement of the inferior parietal lobule in agency. *NeuroReport* 13, 1975–1978. doi: 10.1097/00001756-200210280-00029

Creswell, J. W., and Plano Clark, V. (2011). "Choosing a mixed methods design," in *Designing and Conducting Mixed Methods Research* (Los Angeles, CA: Sage), 2, 53–106.

Creswell, J. W., and Plano Clark, V. (2007). "Designing and conducting mixed methods research," in *Designing and Conducting Mixed Methods Research* (Thousand Oaks, CA; London; New Delhi: SAGE Publications).

Creswell, J. W. (2003). "Research design: Qualitative, quantitative, and mixed method approaches," in *Research design: Qualitative, quantitative, and mixed method approaches* (2. ed.) (Thousand Oaks, CA: Sage Publications).

Creswell, J. W., Plano Clark, V., Gutmann, V., and Hanson, W. (2003). "Advanced mixed methods research designs." in *Handbook of Mixed Methods in Social and Behavioral Research,* eds A. Tashakkori and C. Teddlie. (Thousand Oaks, CA: Sage).

De Bruin, L., Van Elk, M., and Newen, A. (2012). Reconceptualizing second-person interaction. *Front. Hum. Neurosci.* 6:151. doi: 10.3389/fnhum.2012.00151

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

Denscombe, M. (2008). Communities of practice: a research paradigm for the mixed methods approach. *J. Mixed Methods Res.* 2, 270–283. doi: 10.1177/1558689808316807

Depraz, N. (2018). "Surprise, valence, emotion: the multivectorial integrative cardio-phenomenology of surprise," in *Surprise: An Emotion?*, eds N. Depraz, A. J. Steinbock (Cham: Springer International Publishing), 23–52. doi: 10.1007/978-3-319-98657-9_2

Depraz, N. E., Varela, F. J., and Vermersch, P. E. (2003). *On Becoming Aware: A Pragmatics of Experiencing*. Amsterdam; Philadelphia, PA: John Benjamins Publishing Company. doi: 10.1075/aicr.43

Fantasia, V., De Jaegher, H., and Fasulo, A. (2014). We can work it out: an enactive look at cooperation. *Front. Psychol.* 5:874. doi: 10.3389/fpsyg.2014.00874

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

Fenemore, A. (2003). On being moved by performance. *Performance Res.* 8, 107–114. doi: 10.1080/13528165.2003.10871975

Fetters, M. D., Curry, L. A., and Creswell, J. W. (2013). Achieving integration in mixed methods designs—principles and practices. *Health Serv. Res.* 48, 2134–2156. doi: 10.1111/1475-6773.12117

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

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

Fuchs, T., Messas, G. P., and Stanghellini, G. (2019). More than just description: phenomenology and psychotherapy. *Psychopathology* 52, 63–66. doi: 10.1159/000502266

Fuchs, T. E., and Schlimme, J., E. (2009). Embodiment and psychopathology: a phenomenological perspective. *Curr. Opin. Psychiatry* 22, 570–575. doi: 10.1097/YCO.0b013e3283318e5c

Galbusera, L., Finn, M. T., and Fuchs, T. (2018). Interactional synchrony and negative symptoms: an outcome study of body-oriented psychotherapy for schizophrenia. *Psychotherapy Res.* 28, 457–469. doi: 10.1080/10503307.2016.1216624

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

Gallagher, S. (1997). Mutual enlightenment: recent phenomenology in cognitive science. *J. Conscious. Stud.* 4, 195–214.

Gallagher, S., and Zahavi, D. (2008). *The Phenomenological Mind*. London; New York, NY: Routledge.

Gibson, J.J. (1979). *The Ecological Approach to Visual Perception*. New York and London: Psychology Press.

Greene, J. C. (2007). "Mixing methods in social inquiry," in *Mixing Methods in Social Inquiry* (1. ed.), (San Francisco, CA: Jossey-Bass).

Hemakom, A., Goverdovsky, V., Aufegger, L., and Mandic, D. P. (2016). "Quantifying cooperation in choir singing: respiratory and cardiac synchronisation," in *2016 IEEE International Conference on Acoustics, Speech and Signal Processing* (Shanghai: ICASSP), 719–723.

Høffding, S. (2019). *A Phenomenology of Musical Absorption*. Cham: Palgrave Macmillan.

Høffding, S., and Martiny, K. (2016). Framing a phenomenological interview: what, why and how. *Phenomenol. Cogn. Sci.* 15, 539–564. doi: 10.1007/s11097-015-9433-z

Husserl, E. (1989). *Ideas Pertaining to a Pure Phenomenology and to a Phenomenological Philosophy*. Second book, eds R. Rojcewicz and A. Schuwer, Trans. Dordrecht; Boston, MA; London: Kluwer Academic Publishers.

Johnson, R. B., and Onwuegbuzie, A. J. (2004). Mixed methods research: a research paradigm whose time has come. *Educ. Res.* 33, 14–26. doi: 10.3102/0013189X033007014

Koch, S. C., and Fuchs, T. (2011). Embodied arts therapies. *Arts Psychother,* 38, 276–280. doi: 10.1016/j.aip.2011.08.007

Konvalinka, I., Xygalatas, D., Bulbulia, J., Schjødt, U., Jegind,ø, E.-M., Wallot, S., et al. (2011). Synchronized arousal between performers and related spectators in a fire-walking ritual. *Proc. Natl. Acad. Sci. U.S.A.* 108, 8514–8519. doi: 10.1073/pnas.1016955108

Køster, A. and Fernandez, A. V. (in press). Investigating Modes of Being in the World: An Introduction to Phenomenologically Grounded Qualitative Research. Phenomenology and the Cognitive Sciences.

Le Van Quyen, M., and Petitmengin, C. (2002). Neuronal dynamics and conscious experience: an example of reciprocal causation before epileptic seizures. *Phenomenol. Cogn. Sci.* 1, 169–180. doi: 10.1023/A:1020364003336

Lutz, A. (2002). Toward a neurophenomenology as an account of generative passages: a first empirical case study. *Phenomenol. Cogn. Sci.* 1, 133–167. doi: 10.1023/A:1020320221083

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. doi: 10.1073/pnas.032658199

Marsh, K., Richardson, M., and Schmidt, R. C. (2009). Social connection through joint action and interpersonal coordination. *Top. Cogn. Sci.* 1, 320–339. doi: 10.1111/j.1756-8765.2009.01022.x

Martiny, K. (2015a). *Embodying Investigations of Cerebral Palsy, A Case of Open Cognitive Science (Pre-print)*. Faculty of Humanities, Department of Media, Cognition and Communication, Center for Subjectivity Research.

Martiny, K. (2015b). How to develop a phenomenological model of disability. *Med. Health Care Philosophy* 18, 553–565. doi: 10.1007/s11019-015-9625-x

Martiny, K. (2017). *Varela's Radical Proposal: How to Embody and Open Up Cognitive Science*. Constructivist Foundations, 13, 59–67.

Martiny, K., Pedersen, D. B., and Birkegaard, A. (2016). Open media science. *J. Sci. Commun.* 15:A02. doi: 10.22323/2.15060202

Merleau-Ponty, M. (2012). *Phenomenology of Perception*. London; New York, NY: Routledge.

Morgan, D. L. (2007). Paradigms lost and pragmatism regained: methodological implications of combining qualitative and quantitative methods. *J. Mix. Methods Res.* 1, 48–76. doi: 10.1177/2345678906292462

Müller, V., Delius, J. A. M., and Lindenberger, U. (2018). Complex networks emerging during choir singing. *Ann. N. Y. Acad. Sci.* 1431, 85–101. doi: 10.1111/nyas.13940

Müller, V., and Lindenberger, U. (2011). Cardiac and respiratory patterns synchronize between persons during choir singing. *PLoS ONE* 6:e24893. doi: 10.1371/journal.pone.0024893

Nicholson, H. (2005). *Applied Drama: The Gift of Theatre*. Basingstoke; New York, NY: Palgrave Macmillan.

Nicholson, H. (2011). *Theatre, Education and Performance: The Map and the Story*. New York, NY: Palgrave Macmillan.

Parnas, J., and Henriksen, M. (2014). Disordered self in the schizophrenia spectrum: a clinical and research perspective. *Harv. Rev. Psychiatry* 22, 251–265. doi: 10.1097/HRP.0000000000000040

Parnas, J., Møller, P., Kircher, T., Thalbitzer, J., Jansson, L., Handest, P., et al. (2005). EASE: examination of anomalous self-experience. *Psychopathology* 38, 236–258. doi: 10.1159/000088441

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., Baulac, M., and Navarro, V. (2006). Seizure anticipation: are neurophenomenological approaches able to detect preictal symptoms? *Epilepsy Behav.* 9, 298–306. doi: 10.1016/j.yebeh.2006.05.013

Petitmengin, C., and Bitbol, M. (2009). Listening from within. *J. Conscious. Stud.* 16, 10–12.

Petitmengin, C., Navarro, V., and Le Van Quyen, M. (2007). Anticipating seizure: pre-reflective experience at the center of neuro-phenomenology. *Conscious. Cogn.* 16, 746–764. doi: 10.1016/j.concog.2007.05.006

Petitmengin, C., Remillieux, A., Cahour, B., and Carter-Thomas, S. (2013) A gap in Nisbett and Wilson's findings? A first-person access to our cognitive processes. *Conscious. Cognition* 22, 654–669. doi: 10.1016/j.concog.2013.02.004

Petitmengin, C., Remillieux, A., and Valenzuela-Moguillansky, C. (2019). Discovering the structures of lived experience. *Phenomenol. Cogn. Sci.* 18, 691–730. doi: 10.1007/s11097-018-9597-4

Petitot, J., Varela, F. J., Pachoud, B., and Roy, J.-M. (eds.) (1999). *Naturalizing Phenomenology Stanford, Calif*. Stanford, CA: Stanford University Press.

Rallis, S. F., and Rossman, G. B. (2003). "Mixed methods in evaluation contexts," in *Handbook of Mixed Methods in Social and Behavioral Research*, eds A. Tashakkori and C. Teddlie (Thousand Oaks, CA: Sage).

Ruby, P., and Decety, J. (2001). Effect of subjective perspective taking during simulation of action: a PET investigation of agency. *Nat. Neurosci.* 4, 546–550. doi: 10.1038/87510

Sass, L., Pienkos, E., Skodlar, B., Stanghellini, G., Fuchs, T., Parnas, J., et al. (2017). EAWE: examination of anomalous world experience. *Psychopathology* 50, 10–54. doi: 10.1159/000454928

Satne, G., and Roepstorff, A. (2015). Introduction: from interacting agents to engaging persons. *J. Conscious. Stud.* 22, 9–23.

Schilbach, L., Timmermans, B., Reddy, V., Costall, A., Bente, G., Schlicht, T., et al. (2013). Toward a second-person neuroscience. *Behav. Brain Sci.* 36, 393–414. doi: 10.1017/S0140525X12000660

Schmidt, P. (2018). The Relevance of Explanatory First-Person Approaches (EFPA) for understanding psychopathological phenomena. the role of phenomenology. *Front. Psychol.* 9:694. doi: 10.3389/fpsyg.2018.00694

Shaughnessy, N. (2012). *Applying Performance: Live Art, Socially Engaged Theatre and Affective Practice*. London: Palgrave Macmillan. doi: 10.1057/9781137033642

Shepherd, S. (2006). *Theatre, Body and Pleasure*. London; New York, NY: Routledge.

Small, M. (2011). How to conduct a mixed methods study: recent trends in a rapidly growing literature. *Annu. Rev. Sociol.* 37, 57–86. doi: 10.1146/annurev.soc.012809.102657

Swarbrick, D., Bosnyak, D., Livingstone, S., Bansal, J., Marsh-Rollo, S., Woolhouse, M., et al. (2019). How live music moves us: head movement differences in audiences to live versus recorded music. *Front. Psychol.* 9:2682. doi: 10.3389/fpsyg.2018.02682

Tashakkori, A., and Creswell, J. W. (2007). The new era of mixed methods. *J. Mix. Methods Res.* 1, 3–7. doi: 10.1177/2345678906293042

Tashakkori, A., and Teddlie, C. (2008). "Quality of inferences in mixed methods research: calling for an integrative framework," in *Advances in Mixed Methods Research: Theories and Applications*, ed M. *Bergman* (London: Sage Publications), 101–119.

Teddlie, C., and Tashakkori, A. (2009). *Foundations of Mixed Methods Research, Thousand* Oaks, CA: Sage Publications.

Thompson, J. (2009). *Performance Affects*. London: Palgrave Macmillan.

Toro, J. (2020): *Physical Disabilities: An Enactive Exploration of Normal Embodiment (Pre-print)*. Faculty of Humanities, Department of Media, Cognition and Communication, Center for Subjectivity Research.

Toro, J., and Martiny, K. (2020). New perspectives on person-centered care: an affordance-based account. *Med. Health Care Philosophy* 23, 631–644. doi: 10.1007/s11019-020-09977-w

Upham, F. (2018) Detecting the Adaptation of Listeners' Respiration to Heard Music. PhD dissertation. New York University. Available online at: https://pqdtopen.proquest.com/doc/2128010509.html?FMT=ABS

Upham, F., and Mcadams, S. (2018). Activity analysis and coordination in continuous responses to music. *Music Percept.* 35, 253–294. doi: 10.1525/mp.2018.35.3.253

Valenzuela-Moguillansky, C., O'Regan, J. K., and Petitmengin, C. (2013) Exploring the subjective experience of the "rubber hand" illusion. *Front. Hum. Neurosci.* 7:659. doi: 10.3389/fnhum.2013.00659

Varela, F. (1996). Neurophenomenology. A methodological remedy for the hard problem of consciousness. *J. Conscious. Stu.* 3, 330–349.

Varela, F., and Shear, J. (1999). "The view from within: First-person approaches to the study of consciousness," in *The View From Within: First-Person Approaches to the Study of Consciousness* (Thorverton: Imprint Academic).

Venkatesh, V., Brown, S., and Bala, H. (2013). Bridging the qualitative-quantitative divide: guidelines for conducting mixed methods research in information systems. *MIS Quarterly* 37, 21–54. doi: 10.25300/MISQ/2013/37.1.02

Vermersch, P. (1994). *L'entretien D'explicitation (ESF* Éditeur).

Walton, A., Richardson, M. J., Langland-Hassan, P., and Chemero, A. (2015). Improvisation and the selforganization of multiple musical bodies. *Front. Psychol.* 6:313. doi: 10.3389/fpsyg.2015.00313

Walton, A., Washburn, A., Chemero, A., Langland-Hassan,., P., Kloos, H., et al. (2018) Creating time: social collaboration in music improvisation. *Top. Cogn. Sci.* 10, 95–119. doi: 10.1111/tops.12306

Zahavi, D. (2019a). Applied phenomenology: why it is safe to ignore the epoché. *Continental Philosophy Rev.* 1–15. doi: 10.1007/s11007-019-09463-y

Zahavi, D. (2004). Phenomenology and the project of naturalization. *Phenomenol. Cogn. Sci.* 3, 331–347. doi: 10.1023/B:PHEN.0000048935.94012.4e

Zahavi, D. (2017). *Husserl's Legacy*. Oxford: Oxford University Press. doi: 10.1093/oso/9780199684830.001.0001

Zahavi, D. (2019b). Getting it quite wrong: van manen and smith on phenomenology. *Qual. Health Res.* 29, 900–907. doi: 10.1177/1049732318817547

Zahavi, D., and Martiny, K. (2019). Phenomenology in nursing studies: new perspectives. *Int. J. Nurs. Stud.* 93, 155–162. doi: 10.1016/j.ijnurstu.2019.01.014
Zahavi, D. (2011). Varieties of reflection. *J. Conscious. Stud.* 18, 9–19.

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.

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