edition monographs

- *Multikulturalizam i etnicitet*
  Goran Bašić
- *Social Culture: Reevaluating the Paradigm*
  Ljubomir Hristić
- *Slepe ulice istorije: Elita, (dis)kontinuitet i legitimitet*
  Neven Cvetićanin
- *О парадоксу прогреса*
  Владимир Ментус
- *Sudskomedicinska veštačenja u teoriji i praksi medicinskog prava*
  Hajrija Mujović
- *Nasilje i mi: Mediji o nasilju nad ženama*
  Zorica Mršević
- *U susret regionalnoj depopulaciji u Srbiji*
  Vladimir Nikitović
- *Генетички и други здравствени основи дискриминације на раду*
  Сања Н. Стојковић Златановић
- *Populism, Stabilitocracy and Multiculturalism*
  Zoran Lutovac
- *Pravni aspekti nasilja nad starijim osobama*
  Marta Sjeničić
- *Демографска анализа утицаја здравствене заштите и јавног здравља на трендове смртности становништва Србије*
  Иван Маринковић
- *Tržišna ekonomija i poslovna kultura: Nemačka i Japan*
  Marijana Maksimović
- *Opažanje i čulna spoznaja kod Tome Akvinskog*
  Predrag Milidrag
- *Niske stope rađanja i rodne uloge: teorijski okvir i praktični izazovi*
  Ankica Šobot
- Homo experimentalis – *geografija sociooloških eseja*
  Veselin Mitrović
- *Pravni položaj investicionih društava na tržištu kapitala*
  Ranko Sovilj
- *Biheviоralna ekonomija u funkciji održivog razvoja*
  Predrag Jovanović
- *SELF: From One to Many and Back to None*
  Janko Nešić

9 788670 932746

"The scientific contribution of Nešić's book is primarily reflected in an exhaustive review of phenomenological, panpsychist and embodied approaches to the self, as well as their careful analysis and selection of elements that can form a coherent whole and provide a naturalized theory of the substantive minimal self, inspired above all by phenomenological insights and methods, but also compatible with neuroscientific and psychological results."

Miljana Milojević,
Faculty of Philosophy, University of Belgrade

"The book's main value is that it provides a comprehensive comparative analysis and critical review of the most relevant theories of the self, as well as various less known theories, adding significant value to the contemporary debate on the subject."

Vojin Rakić,
Institute of Social Sciences

"Presented in a clear and lucid manner, the book represents a significant scientific contribution that is reflected in a careful analysis of the ontology of the self, attempts to solve the "hard problem of consciousness" within the framework of panpsychism and the theory of integrated information, as well as the metaphysical questions that these theories deal with. Of special importance is Nešić's analysis, which provides a deep insight into the complex aspects of enactivism and its application in the philosophy of mind, theories of the self and psychiatry."

Olivera Vuković,
Faculty of Medicine, University of Belgrade

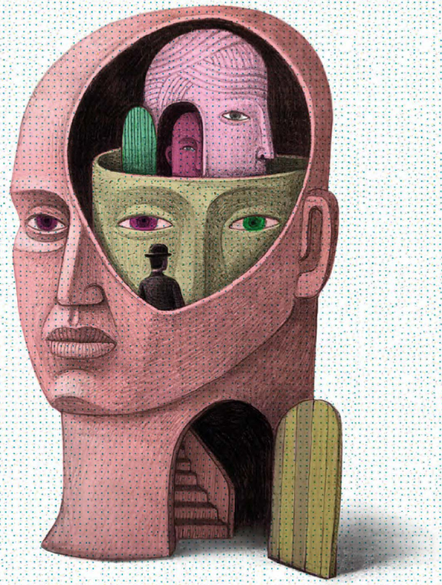
www.idn.org.rs

SELF: FROM ONE TO MANY AND BACK TO NONE
Janko Nešić

Janko Nešić

monographs

SELF: FROM ONE TO MANY AND BACK TO NONE



INSTITUTE OF SOCIAL SCIENCES
Institute of national significance for the Republic of Serbia

About the author

Janko Nešić

Janko Nešić works as a Research Associate at the Institute of Social Sciences, Belgrade. The author received his doctorate in Philosophy from the Faculty of Philosophy, University of Belgrade, with a dissertation titled "Mental Causation Problem: Dispositional Theory Solution". He was a visiting postdoctoral scholar at the University of Fribourg, University of Stockholm and University Clinic of Heidelberg. As part of the project team on the *Sciences of the Origin* project (funded by the Templeton Foundation), he explored philosophical and methodological foundations of the scientific quest for the origins of the universe, life, and mind. His research is on topics at the intersection of philosophy of mind, phenomenology, and philosophy of psychiatry. The author currently focuses on studying pathologies of intersubjectivity in autism spectrum disorder and building an integrative approach to autism that combines phenomenological, enactive, ecological, and neuroscientific perspectives.

SELF:
FROM ONE TO MANY AND BACK TO NONE

Series
Monographs

Series Editor
Predrag Milidrag, PhD

Published by
Institute of Social Sciences
Belgrade, 2024

Publisher
Goran Bašić, PhD

Reviewers
Prof. Vojin Rakić, PhD
Prof. Miljana Milojević, PhD
Prof. Olivera Vuković, PhD

ISBN 978-86-7093-274-6

This monograph was written as a part of the 2024 Research Program of the Institute of Social Sciences with the support by the Ministry of Science, Technological Development and Innovation of the Republic of Serbia.

Janko Nešić

SELF:
FROM ONE TO MANY AND BACK TO NONE

monographs

INSTITUTE
OF SOCIAL SCIENCES
Institute of national significance
for the Republic of Serbia

For my mother Biljana, and my aunt Ninuška,
my closest selves

Contents

INTRODUCTION

I had my first electroencephalogram done when I was eight years old. This uncomfortable ritual involved scrubbing the scalp with alcohol at specific points and attaching electrodes to those points to record the brain activity. After the preparation, the machine started buzzing, an endless sheet of paper was fed in, and the inked waveform patterns emerged. A cryptic chronicle of what I was thinking and feeling, of *Me*, being written in real-time for everyone to see. Could the doctors observe what I was thinking about them and their tedious procedure? Well, even at that age, I did know that they could not see my thoughts, though they could observe if I was drowsing off, that is, if I was not doing what they had instructed me to do, and this made me anxious and scared. Afterwards, looking at this thick book of wavy etchings, which depicted only half an hour of my conscious life, I was not able to connect it with myself; how was this about *Me*? What did these waves on paper say about me and my experiences? The disconnection between what is inside and what outside, the first and third-person perspectives, still haunts me today as it did when I was a child, although I did not refer to it as such.

Later in my life, I enjoyed solitary strolls in our house's backyard in twilight. I gazed with heightened awareness upon its mystical bluish glow, observing the alchemy it performed on the hues of plants, rocks, and soil. In this ephemeral hour, the green of the vegetation took centre stage, a "pure green" that whispered of ancient sensory tales, embracing me in one of humanity's oldest and most universal experiences. Yet, these recollections possess a different essence, a unique quality that exists in the realm of remembrance. These experiences, like threads woven into the tapestry of my being, constitute the very core of who I am.

Sometimes, as I contemplate myself, it is as if these moments have sprouted a life of their own. They carry me, the "pure

I," with them, yet they are distinct from this "I." Upon closer examination, it dawns on me that in each of these moments, past or present, I have been there as the same "I." Though elusive in its definition, this "I" persists, an ever-present companion through the tapestry of time. In my imagination, this primitive self or ego seems to be akin to what the poet and philosopher John Cowper Powys called the *ichthyosaurus-ego* in *The Defence of Sensuality*. This is the name he gives to the "mingling in our 'I am I' of these sub-human and super-human elements", thus baptized to stress the "vegetable-reptile-saurian background of the human soul" (Powys, 1930: 9)

And although I am aware as an adult human being that I have gone through many, even revolutionary, changes in my person (so-called *narrative self*), I still somehow feel like the same subject/self in some sense, besides being the same organism I was. This feeling of continuity is there even though, at times, I do not even recognize myself in the pictures that come from the archaic times of my life, my childhood.

What if I am wrong, and there is no "me" to speak of? What if it just seems to me that there is something *inside*? Is it the same *Me* that was in a childhood photo years ago? What would this *Me* be? Not everyone shares my intuitions about the self; not everyone finds this basic, primordial "me" or "I" in themselves. Moreover, even if I am right about the phenomenological datum that there is something in me, like an I, in metaphysical terms, does it have any substance, or should it be considered a construct, an illusion?

Self is a multifaceted phenomenon. Moreover, there are different kinds of self-awareness and self-consciousness, the awareness of oneself, and these can change in human beings in all sorts of scary ways. Psychopathology can be very illustrative of what the self can be or become. Clinical cases of psychopathological changes to the self are very frightening and mind-boggling. Such cases spark the philosophical questions and debates about the nature of the self, varieties of selves and types of self-consciousness. The benefit is mutual; case studies and research in psychiatry and neuroscience can illuminate philosophical theories of the self and vice versa; concepts and ideas from philosophy

help get to a fine-grained understanding of the workings of consciousness and the self. We give thought and attention (reflection) to something so mundane as consciousness and subjectivity only when it goes awry and when something is wrong, when something changes.

Memory and narrative play an important part in our personal identity in the persistence of the self. How can one speak of any self at a certain point when there are such devastating memory losses in conditions like dementia or Korsakoff syndrome, as in the famous case of Jimmy G., whose story was delicately told by Oliver Sacks in his book *The Man Who Mistook His Wife For a Hat*. Jimmy could retain memory for only a couple of seconds and had dense amnesia that went all the way to 1945 (he was admitted in 1975). Still, Sacks argues that even in the face of such a neurological catastrophe, something was preserved, and Jimmy was able to "find himself" in spirituality and the church. Another fascinating example of anomalous bodily self-awareness comes in the very next chapter of that book — the story of Christina, *The Disembodied Lady*, whose body suddenly became blind, unable to "see" itself, by loss of proprioception, as she began to not feel her body as her own (experiencing the loss of body-ego).[1] And what should we make of the split-brain patients and craniopagus twins? How many subjects are there in these strange cases?

A significant portion of this monograph will be dedicated to the phenomenology of subjectivity, of what it feels like to be a subject of experience, and to various self-awarenesses that make up the feeling of ourselves, our *self-experience*. Modern phenomenological and neuroscientific exploration has shown that in certain states, like meditation and psychedelic experiences, as well as in many pathologies, our self-experience can radically change. In these instances, the structure of self-awareness seems to break down, and the "ego" can dissipate and dissolve (or so are such subjects aware!). There is talk of *ego dissolution, disintegration* or *ego death*. In mystical experiences, the boundary between the self and the world sometimes weakens and disappears.

[1] Somehow, she managed to compensate this loss with the aid of strengthening the visual body-image.

In these states, radical changes are evident both from the inside and the outside, subjectively and objectively (loss of some form of self-awareness and loss of typical coordination of the default mode network of the brain). How do we make sense of these experiences, and do they tell us something about the self?

There are different types of selves, or layers within the self that are usually distinguished in the ever-growing literature on the subject (pre-reflective, reflective, minimal, bodily, narrative, etc.). However, here I will mostly be concerned with what makes the minimal kind of self, the subject of experience, and to try and see if there is such a thing. I am certain that it indeed exists, but it is not what we have thought it would be, to say the least. Still, theoretical considerations under discussion here bear on higher levels of the self (narrative). In this book, I plan to survey and examine, sometimes in detail, various modern metaphysical theories of self. Although I will aim to cover all the main aspirants for THE philosophical theory of the self, I must confess that this will still be a very subjective point of view on the debate.

My choice of theories to which I will dedicate the most attention will be to include examination of approaches that have not been given enough (selfless) room in the literature (e.g. modern versions of dualism, panpsychism, structuralism, enactivism, predictive processing). As I tell the story of the Self, I will simultaneously provide some possible arguments of my own for those positions, putting forward suggestions on how to solve certain theoretical problems and even proposals for new directions of inquiry and what I believe to be the future ways towards a deeper understanding of selfhood. Despite some wild ideas and positions presented in the book, I will strive towards naturalism about selfhood. The overarching (implicit) theme of the monograph will be the *naturalisation of phenomenology*.

In the first Chapter, I lay out a slightly revised traditional division between different approaches to the nature of the self that I think will be useful to steer us along the selfish path. The remainder of Chapter 1 is dedicated to the phenomenology of subjectivity and the self — to what it feels like from the inside, subjectively, to be someone, a subject of experience. This is the most interesting perspective on the self, and it must be examined

in as much detail as possible. I think that subjective data of consciousness will always play an important role in understanding the self, be it from the neuroscientific or philosophical point of view. I will use the phenomenological data from this chapter and, in the next part of the book (Chapter 2), present a possible argument for the substantivalist self. In the third Chapter, I return to some less-known minimal self-theories coming from the Phenomenological tradition.

In Chapter 4, I survey panpsychist views on the subject of experience and try to provide answers to some famous combination problems that plague these metaphysical theories (the subject summing problem being the hardest and most interesting for the topic of this book). Many of those solutions will be controversial, though based on the findings and arguments of the previous chapters, and therefore providing at least plausible ideas for panpsychists to think about in the future. I will endeavour to highlight that the issue of subjectivity and subjects of experience, although thoroughly discussed in the Phenomenological tradition and analytical metaphysics of today, has not been properly thought out by the panpsychists. Especially now, when all forms of panpsychism have been given a metaphysical comeback, the problem of subjectivity and self needs to be dealt with care.

Continuing in that direction, along the panpsychist lines, I will also problematize subjectivity in Integrated Information Theory (IIT) in Chapter 5. Having introduced Integrated Information in Chapter 6, I will discuss the structuralist interpretation of the self — the structural realist theory of the self. There, I plan to propose some alterations to the existing structuralist theory of the self and defend the view that such a theory of the self and consciousness is more aptly combined with the neutral monist ontology.

Finally, Chapter 7 delves into the most recent enactivist and predictive processing theories of the self and subjectivity. I will be specifically interested in the ecological-enactive (EE) interpretation of the free energy principle (FEP) and predictive processing (PP) when it comes to self-models. Thus, I will be going a full circle, through the “thick” and “thin” of selfhood.

Chapters 1, 2 and 3 are based on parts of my paper “I Am Mine: From Phenomenology of Self-awareness to Metaphysics of Selfhood” (Nešić, 2023b).

Chapter 4 is based on my articles “Against Deflation of the Subject” (Nešić, 2017) and “Subject Emergence in Panpsychism” (Nešić, 2018b).

Chapter 5 is based on my “Does Integrated Information Lack Subjectivity?” (Nešić, 2018a).

Chapter 6 is based on my “Towards a Neutral-Structuralist Theory of Consciousness and Selfhood” (Nešić, 2022a). I am grateful to Taylor & Francis for allowing me to use this material.

Chapter 7 is based on parts of my “Enactive Framework for Psychiatric Disorders” (Nešić, 2022b) and “Ecological-enactive Account of Autism Spectrum Disorder” (Nešić, 2023a). I am grateful to Springer Nature for allowing me to use this material.

1. PHENOMENOLOGY OF SUBJECTIVITY

> My hands went up to my face. My face was there from the inside. My body was more than just a series of textures that my hands knew, an image my eyes knew, a series of sounds my ears knew, and a pattern of movements. I cried out in a desperate whisper, "Oh my God, I've got a body."
>
> *Somebody Somewhere* (Williams, 1994: 262)

1.1 THEORIES OF THE SELF

We cannot begin this journey into the heart of the Self without the notion of what theories of the self are on the metaphysical table. There are two traditional camps that stand out: the *substance theory* and the *bundle theory* of the self. I will add a third, a more recent position — the minimal/phenomenal self. We can define these three main categories of metaphysical theories about experiencing subjects (selves) in the following manner. These categories are based on how we want to understand the self/subject in relation to experience. Are they just experiences in relation, something substantially more than just related experience, or do they somehow make a whole? In addition, the self/subject could be viewed, in terms of its nature, as physical, purely mental, mental and physical, or as something neutral.

Let us start with the bundle theory — if the subject is understood as a kind of a bundle, then a subject is individuated by experiences (identity conditions of persons are specified in terms of relations between mental states), which are psychological modes — the subject is just experiences and their relations; selves are collections of properties. (e.g., Hume, 1978; Parfit, 1984; Dainton, 2008). Selves reduce to experience.

In case the self is a kind of "minimal subject", it is identical to the *subjectivity* (with *a feature*) of experience. For example, Dan Zahavi (Zahavi, 2014) defends *experiential minimalism* in which the *for-me-ness* or *first-personal givenness* dimension of phenomenal consciousness *is* considered the *minimal experiential self*. He is explicit on this: "More precisely, the claim is that the (minimal or core) self possesses experiential reality and that it can be identified with the ubiquitous first-personal character of the experiential phenomena" (Zahavi, 2014: 18). This modern approach is a descendant of theories from Phenomenological tradition. The "minimal/thin" theories are deflationist about the self and are similar to the bundle view.

Galen Strawson's (Strawson, 2009) *sesmet* theory of self (*sesmet* being the acronym for "subject-of-experience-as-single-mental-thing") could also be called *thin self theory*, with the difference being that the experiential self is *diachronically* persistent in Zahavi's account. Guillot (Guillot, 2017: section 3.5.2) discusses minimalism about the self: "the self, or at least a form of selfhood (the "minimal self" or "core self"), is identical either with experience or with some part or intrinsic property of experience", citing Zahavi (Zahavi, 2005) and Williford (Williford, 2015) as proponents while saying that Strawson's view "bears a resemblance to this type of minimalism". The idea of Strawson's s*esmet* seems to be peculiar. Strawson holds that selves are objects (Strawson, 2009: 298) because of their "strong unity", though he is distancing himself from the traditional notion of an individual substance. With every experience, there has to be a subject of experience. Experiencing involves a subject, and this is both a metaphysical and a phenomenological claim. Sometimes, Strawson claims that there is an identity between the subject and experience. In a sense, Strawson could be understood as claiming that selves are substances (in a new, more relaxed sense of *substance/object*), though not enduring substances. If this is the case, his theory would belong in the first category. I do not have the space here to analyse Strawson's view in more detail.

What do we mean when we say that the subject of experience is a kind of a *substance*? The claim is that an experiential subject is a kind of metaphysical entity that is a bearer of experiential

properties, upon which experiential properties are instantiated, and is not in itself a kind of property (e.g., P.F. Strawson, 1959; Chisholm, 1969; Lowe, 1996; Nida-Rümelin, 2018). The subject is that which endures among many changing experiences. Hence, the instantiations of experiential properties in subjects are types of events, that is, *experiences*. For the subject involved, there is a kind of phenomenology involved, it is like something to undergo the experience because experiential properties are such that there is something it is like to have them. Thus, if the self or subject is a substance, it metaphysically unifies and individuates experiences (experiential properties) as their bearer. Lowe refers to (Lowe, 1996: 9): "selves or persons as *substances* — that is, as enduring bearers of successive states and in no way reducible to mere successions of those states". Experiencing subjects, in that case, are fundamental.

One could add the fourth category of *no-self* accounts, those that contend that the self simply does not exist. No-selfers are usually denying the substantiality and individuality of the subject.[2] Metzinger (Metzinger, 2003, 2011) is a famous proponent of the *no-self* theory, though he acknowledges that there is a *phenomenology of substantiality*. If there is relevant self-awareness, it would provide us "with a direct and epistemologically relevant form of *acquaintance* with ourselves" (Metzinger, 2011: 284), and we could grasp our own nature. Still, he is not of the opinion that phenomenology determines metaphysics. Metzinger thinks there are two strategies to argue for ontological anti-realism about the self — one is to discuss the "phenomenon of self-consciousness and related folk-phenomenological discourse, including its unwarranted metaphysical assumptions", and the other is to deny that there are any substances at all. One of the options open for the anti-realist is the bundle theory (self as a mere collection of properties). Some of the no-self accounts are found within the bundle category, and others are close to minimal views (Buddhist and enactive views, for example).

[2] Read Zahavi, 2014: Ch. 4.1 for discussion.

As this book progresses, we shall see how matters of the self become more complex and how modern theories do not easily fit into these categories. For now, this will suffice.

1.2 SELF-AWARENESS

We can now turn to phenomenology[3] in the endeavour to properly understand what has broadly been called *the subjective character* of consciousness (and other key notions found in these debates, like those of *self-awareness* and *mineness*). Since the self is tightly connected with the notion of self-awareness, it is only natural to start the journey with the phenomenological datums that, perhaps, give us a reason to believe in the reality (or illusion) of the self.

The concept of subjectivity or subjective character of consciousness has been underappreciated in modern analytic theories of consciousness. Others have argued for the project of phenomenological contribution to the philosophy of mind and to the general mind/body debate: "Philosophical phenomenology can offer much more to contemporary consciousness research than a simple compilation of introspective evidence" (Zahavi, 2005: 5). Phenomenology thus construed is not just introspective analysis of experience. Subjectivity has always been one of the main interests of phenomenological investigations, and it is only natural to use such theories when trying to understand the problem of the nature of subjects in panpsychism. Phenomenology could help us understand what a subject of experience is, and only then could we hope to resolve the unity of consciousness, the boundary problem and the subject-summing problem of panpsychism. It would be beneficial if the contemporary debate on deflationism in panpsychism would be more thoroughly informed by the phenomenological concepts of subjectivity/mineness/pre-reflective

[3] Phenomenology can refer to both the subjective character of consciousness, and the Phenomenological tradition, and I will try to make it clear in the text which of these uses I have in mind. These are not unrelated since the notions that describe what can be found in the subjectivity of consciousness mostly come from the Phenomenological tradition.

self-awareness. Specifically, the problem of consciousness' unity will be addressed, and a different solution will be based on the first-person givenness account offered. I will discuss how the inclusion of pre-reflective self-consciousness affects these matters.

There are two distinct aspects of phenomenal conscious states: *something it is like to be* in a mental state and *what it is like to be* in that state. Phenomenally conscious mental states have qualitative and subjective characteristics (Levine, 2001; Kriegel, 2009). These are separate questions of the subjectivity and quality of consciousness, yet these very often seem to be lumped together. The subjective designates how a certain mental state feels to the subject. Upon a more precise reading, it is revealed to be actually meaning that the subject is somehow present in experience as a kind of self-awareness. Mineness, subjective character, and for-me-ness are sometimes meant to reflect or stand in for the *pre-reflective self-consciousness* of the Phenomenological tradition. That there seems to exist something like self-consciousness in the pre-reflective and pre-conceptual sense is held as highly plausible by many phenomenologists and philosophers of mind. It is also called *pre-reflective self-awareness*. This self-consciousness in question is not of the cognitive kind, deployed in I-thoughts, but minimal and non-reflexive. Mineness refers to the feeling that all experiences in consciousness are "mine", synchronic and diachronic. However, this term has several readings, and they can be very different. In her recent paper, Marie Guillot proposes that subjective character refers to several distinct notions that are being confused by some authors: *for-me-ness* (a relation of awareness between a subject and an experience), *meness* (a reflexive relation of awareness a subject has to itself) and *mineness* (a relation of awareness between subject and a fact that it owns the experience) and all these refer to relations of awareness between a subject and its experiences (Guillot, 2017: 32). The third notion is the strongest, but all three imply the presence of a subject of experience.

This means that, besides the regular content of consciousness, there could be something more, the feeling of being a self or a subject. Self-consciousness can be understood in many ways, but we need the fundamental type, *pre-reflective self-awareness*,

the best candidate for the type of awareness that provides a grasp of the nature of the experiencing subject. When we arrive at a clearer understanding of pre-reflective self-awareness, we need to show that acquaintance and self-acquaintance are possible and, thus, a way to the nature of the experiencing subject.

Pre-reflective self-consciousness/awareness, mineness, me-ness, for-me-ness — different philosophers conceive many different concepts, and they do not necessarily signify the same phenomenon, but they all circulate in the literature as pertaining to subjectivity that points to a subject or having something to do with a subjective point of view. Mineness and subjective character are sometimes meant to stand for the pre-reflective self-consciousness of the Phenomenological tradition. That there seems to exist something like self-consciousness in the pre-reflective and pre-conceptual sense is held as highly plausible by many phenomenologists and philosophers of mind. This self-awareness is not cognitive, deployed in I-thoughts, but minimal, non-reflexive, which many have defended as pre-reflective self-consciousness. All these concepts are about properties or aspects of consciousness.[4]

1.3 MINENESS

It has become common in contemporary analytical philosophy of mind to hold that consciousness has a *subjective* as well as a *qualitative* aspect, that there is a difference between what an experience is like and its being like something for its subject (e.g., Levine, 2001; Kriegel, 2009).[5] This subjective dimension of consciousness has been understood differently by different

[4] For discussion and criticism of some uses of these terms, see Siewert 2013, Nida-Rümelin 2014, Guillot 2016. They have shown what lies behind these notions and how we should work towards developing better concepts that describe our phenomenology more accurately. Among modern philosophers who take seriously such notions of subjectivity are Zahavi (Zahavi, 2005, 2014), Gallagher (Gallagher, 2000), Fasching (Fasching, 2009), Kriegel (Kriegel, 2009), Strawson (Strawson, 2009), Levine (Levine, 2001), Shoemaker (Shoemaker. 1996), to name just a few.

[5] From Nagel (Nagel, 1974) to Zahavi (Zahavi, 2014); two *dimensions* of experience.

philosophers. This side or aspect is sometimes called for-me-ness, me-ness, mineness, first-personal givenness or simply subjectivity.[6] Mineness and subjective character are sometimes meant to stand in for the *pre-reflective self-consciousness* of the Phenomenological tradition. Many phenomenologists and philosophers of mind maintain that something like self-consciousness exists in the pre-reflective and pre-conceptual sense. This feature of consciousness is also called *pre-reflective self-awareness*.[7] The self-awareness in question is not of the cognitive kind, deployed in I-thoughts, but minimal, non-reflexive.

Now, mineness could be a misleading term. Are experiences phenomenally presented as mine? One will not, perhaps, find any feature or property of an experience, a stamp or a mark, that would say that *it is mine*. Even if phenomenal consciousness would have such a feature, does it commit one to the existence of a subject of experience? There is much imprecise talk when the matter of subjectivity is concerned. Some of the confusion behind the use of such notions was cleared by the work of Siewert (Siewert, 2013), Nida-Rümelin (Nida-Rümelin, 2014, 2017) and Guillot (Guillot, 2017). They have shown how we should work towards developing better and more accurate concepts based on our phenomenology.

We must be cautious when using the umbrella term "subjective character" because it can designate things that are essentially very different. Nida-Rümelin (Nida-Rümelin, 2014, 2017) shows there are three interpretations of "subjective character": *basic intentionality, primitive awareness and awareness of basic intentionality*. What she calls *awareness of basic intentionality* is *pre-reflective self-awareness*. Nida-Rümelin argues that awareness of basic intentionality cannot have the structure of basic intentionality and is not in itself experiencing.

---

[6] Also, *inner awareness* (Kriegel, 2009; Farell & McClelland, 2017).

[7] When I say *self*-awareness, I mean *awareness of the self* and not *awareness of awareness* as in the "higher-order" and "self-representational" theory of consciousness. For a great discussion on the relation between phenomenality and self-consciousness, see Siewert, 2013. Pre-reflective self-awareness could be present universally in every conscious creature that is a subject of experience, and consequently in every episode of experiencing.

Philosophers of consciousness from the Phenomenological tradition onward have been arguing that there is something special about so-called pre-reflective self-awareness. Such awareness is very hard to find in your attention because it is ever-present. The case is such that a subject is never an *object* in its own stream of consciousness. It is not like one is turning "the mind's eye" inward. Nida-Rümelin (Nida-Rümelin, 2014) would say that the awareness of basic intentionality (self-awareness) is not in itself an experience that exhibits basic intentionality. The subject is not presented to itself "as an object".[8]

As I said, Guillot (Guillot, 2017) has proposed that *the subjective character* refers to several distinct notions confused by certain authors: *for-me-ness* (a relation of awareness between a subject and its experience), *me-ness* (a reflexive relation of awareness a subject has to itself) and *mineness* (a relation of awareness between a subject and a fact that it owns the experience, the fact of ownership) and all these are about relations of awareness between *a subject* and *its experiences* (Guillot, 2017: 32).[9] These are distinct properties not to be conflated. Guillot supports her tripartite framework using the examples from depersonalization syndrome and thought-insertion cases. Mineness and me-ness are not universally present, though for-me-ness seems to be so, as she concludes.[10] Guillot argues that the property of mineness can be lacking in certain cases, like schizophrenic "thought-insertion".

[8] Apart from Nida-Rümelin (Nida-Rümelin, 2014, 2017), such a "non-objectual" view of pre-reflective self-awareness in modern philosophy of mind is also to be found in Zahavi (Zahavi, 2014). However, it is arguable what does Zahavi exactly imply by "non-objectifying form of self-consciousness". He alternates between *subject-self-consciousness* and *state-self-consciousness*. Siewert contends that the "presence" or "givenness" of experience, i.e. how experience is phenomenally "for me" or "mine", should be understood as a kind of self-awareness: "think about the way the viewpoint of the looker is implicit in how things look" (Siewert, 2013: 31). A form of self-awareness is built into the experience: "marginal" *awareness of oneself as a looker*.

[9] Howell and Thompson find that these notions are about *Phenomenal Me-ness*, for which there are two conditions: *The Phenomenal Condition* — Phenomenal Me-ness has to contribute to a subject's total phenomenal character. *The Representational Condition* — Phenomenal me-ness must in some way present the self (Howell and Thompson, 2016: 4).

[10] What Zahavi has in mind when he talks about *mineness* is actually the first notion, i.e. *for-me-ness*. In Marie Guillot's interpretation of these concepts, *me-ness* is what Nida-Rümelin calls *pre-reflective self-awareness*.

Such patients may lack mineness, that is, awareness of experience ownership, but they could still have self-awareness (in the sense of *me-ness*).[11]

Selfhood is one of the structures of experience studied by phenomenological psychopathologists.[12] Modern phenomenological psychopathology uses novel instruments to examine these disturbances of the self (of *self-awareness*/*self-experience,* to be more exact). EASE and EAWE are the first phenomenological instruments developed to quantitatively measure the severity of phenomenological disorders. The EASE (*Examination of Anomalous Self-Experience*; Parnas et al. 2005) is a list of symptoms used for semi-structured phenomenological investigation of anomalies/disorders of subjective experiences associated with the experience of basic or minimal self-awareness. The scale is intended for use in disorders within the schizophrenia spectrum that are essentially characterized by disturbances of subjectivity. The scale focuses exclusively on abnormal features of the self and self-awareness. The EAWE (*Examination of Abnormal Experiences of the World*; Sass et al., 2017) is also a semi-structured interview instrument. This instrument aims to explore experiences related to different aspects of the lived world and explores six dimensions of experience: 1) space and objects, 2) time and events, 3) other people, 4) language, 5) atmosphere, and 6) existential orientation. Like EASE, EAWE is predominantly associated with disorders within the schizophrenia spectrum.[13]

Disturbances of self-experiences have been in psychiatry for a long time. As Gipps (Gipps, 2022) notes, both Jaspers (Jaspers, 1959/1963) and Schneider (Schneider, 1959) thought that

[11] There are several possible views on the prevalence of these features in consciousness. Farell and McClelland (Farell & McClelland, 2017: 4–5) see three options: *Universalism* (*inner awareness*, as they call it, is present in all non-reflective experiences), *Typicalism* (not present in atypical cases) and *Absentism* (never present). Three forms of inner awareness that they make distinct, following Guillot, are for-me-ness, me-ishness and mineness. In their terms, Nida-Rümelin "is a universalist about for-me-ness. She also seems to be at least a typicalist, and perhaps a universalist about something similar to me-ishness and something similar to mineness" (Farell & McClelland, 2017: 12).

[12] Others being *intersubjectivity, embodiment, affectivity, understanding, temporality, spatiality.* See Nelson et al. 2021.

[13] For details see Jerotić & Nešić 2023.

there is some kind of basic "sense of mineness" that is absent in "passivity experiences" (self-disturbances), and when this mineness is lacking experiences seem to be alien, "not mine".[14] Now, it is not clear what exactly is the mineness that they are referring to, while that there is a lack of proper understanding of this phenomenon is explicitly stated by Schneider (Schneider, 1959: 121, 124; cited in Gipps, 2022: 121).[15] Although there is much talk of the ego and mineness, insufficient conceptual dissection and analysis are given.[16]

[14] It was Schneider who put forward passivity experiences as the "first rank" symptoms of the schizophrenias (Gipps, 2022: 110).

[15] Consult Gipps' (Gipps, 2022) Chapter 5, *The Divided Self,* for a critical discussion on the sense of "sense of mineness". He is sceptical that mineness accompanies all of our experiences and says that modern phenomenological psychopathologists like Parnas, Sass, and Zahavi follow The Heidelberg School in seeing the loss of mineness in the passivity experiences. Gipps thinks that such accounts are "incoherent" and argues for an apophatic approach to the psychopathology of psychotic illness.

[16] In a very recent paper, Fazakas, Bois, and Gozé (Fazakas, Bois, & Gozé, 2023) re-introduce the concept of *coenesthesia* as phenomenological materiality to understand the minimal self and its disturbances. They argue that it will bring with it the bodily *thickness* and density of "for-me-ness". With this in mind, the anomalies of coenesthesia can be found in the EASE.

2. SUBSTANCE OF EXPERIENCE

> Our innermost self, as we grow more and more conscious of it, surprises us again and again by new explosions of feeling drawn from emotional, nervous, and even chemical reactions; but for all its surreptitious dependence on these impulses, its inner report upon its own nature is that it is a clear, hard, enclosed, secretive nucleus with a detached and independent existence of its own. Our reliance upon this introspective report may easily be shaken by logical argument; but it is not often that any argument, however plausible, disposes of the feeling of this interior identity, of the feeling of this integral "I am I," underlying the stream of our impressions.
>
> *The Meaning of Culture* (Powys, 1930: 16)

2.1 CAN WE HAVE OUR SUBSTANCES BACK?

The Cartesian substance view has for a long time been the orthodoxy of the Western metaphysics of the self. In this chapter, I will be discussing some modern variants of the view. For example, E. J. Lowe held "the ownership view" and argued against what are considered to be neo-Humean and neo-Lockean theories of personal identity, though he did this from different grounds, namely from proper grasp of one's self-knowledge. In *Subjects of Experiences,* he concludes that: "The self must be conceived of as having the status of a substance vis-a-vis its thoughts and experiences — they are 'adjectival' upon it (are 'modes' of it, in an earlier terminology), rather than it being related to them rather as a set is to its members." (Lowe, 1996: 195). "Property-instances are ontologically dependent entities, depending for their existence and identity upon the individual substances which they characterize, or to which they 'belong'" (Lowe, 2006: 27).

It can be noted that Lowe finds the doctrine of 'bare particularity', i.e. that there is a 'substratum' or 'bare particular' supporting property-instances, indefensible. He maintains that the modes belong to the individual substance itself. "I contend that modes are 'particular ways objects are', and as such are ontologically dependent upon objects in a much stronger sense than, according to a trope theorist, any trope can be ontologically dependent upon other tropes in a bundle of compresent tropes" (Lowe, 2006: 97).[17]

In Lowe's theory of interactive dualism (Lowe, 1996, 2000, 2008, 2013), the mental does not start any new physical causal chains; a mental event causes a physical fact. Mental event M causes it "to be the case that certain physical events, P1, P2, ... Pn, have a certain physical effect, P" (Lowe, 2008: 54). Mental event M is not a direct cause of P1 or any other physical event but of the whole causal chain. What it brings is a *state of affairs*. Therefore, there is a distinction being made between event causation and fact causation. Mental events make the causal tree of neural events converge on a bodily movement non-coincidentally, claims Lowe. Without the mental influence, such a convergence would be a thing of chance. The mental is also invisible to the scientist researching the physical events leading to the movement. Mental does not control the electrochemical signals like tubes are directing the fluid's movement. There is no redistribution of energy or moment because a dualist does not endorse a transference theory of causation.[18]

The causal relata found in Lowe's account are *individual substances*. These are concrete bearers of properties (and so, of powers). They are ontologically independent and never causally inert. In Lowe's view, all causation can be understood as, fundamentally, substance causation. Only substances possess powers; any talk of events or properties possessing powers is just derivative. The mental will, Lowe concludes, should not be alienated from the agent it belongs to. Agent does not need some power

[17] See Lowe 1998, 2006 for more on substances and bundles.

[18] I go into much more detail about Lowe's theory of mental causation in Nešić, 2015.

over his will; it is already the power of the agent, agent simply exercises this will (Lowe, 2013: 157–167).[19]

When it comes to answering the question of the nature of selfhood, most contemporary philosophers of mind have gravitated towards deflationary and eliminativist accounts.[20]

In this Chapter, I will propose one possible phenomenological argument for the claim that the self is a substance. From the phenomenology[21] of pre-reflective self-awareness we can argue for the thesis that the subject of experience is a kind of substance — that the phenomenological situation points to the metaphysical structure of our experience in favour of the substance view.

If a subject is pre-reflectively aware of being the one individual who has experiences (it is revealed to the subject that it is the bearer of experiences, their unifier and individuator), then this would show that the subject is, indeed, a substance. I will try to discuss in the course of the chapter whether all the conditions for this claim could be satisfied. I also intend to argue that the debates on the phenomenology of pre-reflective self-awareness and the metaphysics of the self are closely related.

2.2 I AM MINE: SELF-ACQUAINTANCE

If there is self-awareness and this awareness can afford self-acquaintance, we are on our way to better understand the metaphysical nature of the self. Still, it must be shown that acquaintance with oneself (self-acquaintance) is possible. Modern acquaintance theory comes from Russell and denotes the closeness and intimacy of the subject to its experiences (to the experiential properties of consciousness). The consequence of this

[19] More on Lowe's preferred version of dualism in Nešić 2013a, 2013b.

[20] Even those who are sympathetic to non-physicalist theories of consciousness are still deflationists about selves and subjectivity. See Strawson, 2009 and Chalmers, 2015 for similar claims.

[21] Here, I do not refer to the Phenomenological tradition and philosophical method but to the phenomenality of experience, the "what-it's-likeness" of experiences (and the subjectivity of experience), though many crucial ideas on the nature of pre-reflective self-awareness come from phenomenologists like Husserl, Merleau-Ponty and Michel Henry.

closeness is that the nature of experiences is *revealed* in a certain way to the subject involved. The most likely candidates for the things we are directly aware of or acquainted with in this way are our *experiences*, the experiential properties of consciousness. Other candidates are the *subjects* of those experiences.[22] Modern proponents of the acquaintance approach to introspective or phenomenal knowledge are Gertler, Goff, Horgan and Kriegel, and Nida-Rümelin (Gertler, 2012; Goff, 2015, 2017; Horgan & Kriegel, 2007; Nida-Rümelin, 2007, 2016).

Brie Gertler's approach is explicitly defended as more "modest" than Russell's theory, though it is a descendant of this theory (Gertler, 2012). Gertler's *acquaintance approach to introspective knowledge* can be expressed by the following claim, we can sometimes *directly* grasp our experiences and in such situations, we form phenomenal concepts and introspective judgments about those experiences of ours. These make up our *knowledge by acquaintance*. As Gertler formulates it, the main thesis is that in grasping experiences, phenomenal reality "intersects" with the epistemic (Gertler, 2012: 94–95). Three conditions need to be satisfied for something to be a judgement of introspective knowledge. Those introspective judgements are directly tied to their truthmakers. For their justification, they only depend on the subject's conscious state. They are more justified than empirical judgements (Gertler, 2012: 100). The *Acquaintance* is a claim which involves that the gap between *epistemic appearances* and *phenomenal reality* is sometimes filled. There is also a *metaphysical* claim here since phenomenal reflects the metaphysical reality: judgements are directly tied to their truthmakers — experiences, that is, experiential events.[23]

Acquaintance is the thesis that our intimacy with experience puts us in the position to know the nature of the thing we are acquainted with is revealed to us. Goff calls it the *Real Acquaintance* and defines it as: 'A psychologically normal subject can

[22] Which seems to be left out of the acquaintance discussion in most cases.

[23] Please note that in the Russellian version of acquaintance, the very relation of acquaintance is between a subject and a thing (sense-datum). In modern accounts, it seems to be the relation between an introspective judgment and its truthmaker.

come to know the real nature of one of her phenomenal qualities by attending to that quality' (Goff, 2015: 124). A closely related thesis he proposes is: *'Phenomenal Certainty*: A psychologically normal subject is able to put itself into a situation in which, with respect to one of its phenomenal qualities, it is justified in being certain that that quality is instantiated (where to be certain that P is roughly to believe with a credence of 1 that P)' (Goff, 2015: 124). That is to say, phenomenal knowledge is completely infallible. When a subject has an experience, there can be no doubt that the subject has it from the inside, that the given experiential property is being instantiated (as translated to the *framework of experiential properties* terminology). Goff's thesis of *Phenomenal Certainty*, which is not only implied by *Real Acquaintance*, is encountered again as explained by *Real Acquaintance* because it is a plausible thesis in its own right. In Goff's theory, it is coupled with *Phenomenal Insight*.[24]

If there is acquaintance, then the Revelation thesis is true. To know the nature or essence of a property (phenomenal property P) is to know what it is for the property to be instantiated. If I have a sensation of purpleness, then I know that an experiential property of being phenomenally presented with purple is instantiated in me.

Why is acquaintance important? Because when one is acquainted with something, the *nature of the thing is revealed*. Why wouldn't the same hold true for self-awareness, not just awareness of the experience? Self-awareness is the awareness that the self has of itself that *is direct and immediate*, unmediated (Horgan & Nichols, 2016). So, it would be natural to expect that we are thus acquainted with ourselves and have self-acquaintance. And if this is "real self-acquaintance", then the nature of the self is revealed to us in acquaintance. If the self can satisfy these

[24] "*Phenomenal Insight*: We have rich a priori knowledge concerning our phenomenal qualities." (Goff, 2015: 128). Goff defends "*Phenomenal Transparency*: Phenomenal Transparency is the thesis that phenomenal concepts reveal the essence of the states they denote. According to Revelation, when a person attends to a token conscious state under a direct phenomenal concept, the complete nature of the type to which it belongs is apparent to her; this entails Direct Phenomenal Transparency: the thesis that direct phenomenal concepts are transparent" (Goff, 2017: 108).

requirements, then it can be claimed that we have self-acquaintance in addition to acquaintance with experience (properties).

It seems plausible that for one to argue that the self is a substance and that the self knows this from its experience, one would need a premise that would state the possibility of a subject being *acquainted* with oneself. To know its own nature, a subject must have the proper ability to know this nature to access it. However, one need not expound on ambitious notions of acquaintance in order to do so. I can be wrong about the precise content of some experiences, but I cannot be wrong that I am having some experiences right now, whatever they might be.

One is especially acquainted with oneself because (pre-reflective) self-awareness or self-presence is so intimate that it is immediate and direct (unmediated). Is *self-acquaintance* as plausible as acquaintance with experiences, or do we need additional arguments for it? Many believe that when we are directly aware of something, we are acquainted with it. If we have direct awareness of the self, if there is self-awareness, then the self or subject is acquainted with itself. Duncan has argued that the self passes what he calls *The Doubt Test*, which is a test for acquaintance with something (Duncan, 2015). *The Doubt test* can be found in theories from Descartes' to Russell's, but also in modern theories, like Gertler's (Gertler, 2012) approach and Horgan and Kriegel's (Horgan & Kriegel, 2007). This test states that if we cannot doubt the existence of something being presented to us in awareness, then we are acquainted with it. We can doubt that the object producing my experiences of it exists, but I cannot imagine any sceptical scenario that would make me doubt that I have any experiences in the first place.

Duncan points out that, in the case of an acquaintance with our experiences, we are in the position to be aware of their essence, but it also seems to some philosophers that this is not the case in self-awareness. There could exist an *asymmetry* between experiences and the self. Acquaintance with the self is only partial, revealing only some aspects. But there is no real asymmetry here. The self is as much (directly) revealed as the experience. Both experiences and the self could have hidden aspects unrevealed. What is presented, though, is that being directly aware of.

Though *prima facie,* it may look as if there is a difference between acquaintance with experiences and acquaintance with the self in the sense there is an *appearance/reality gap*, Duncan argues that such is not the case. Experience is just as it seems to be, and the subject's properties could be misleading in the way experience would not be. But the subject/self and the experience are on par with this; the same may be said for experience. There could be no appearance/reality gap with aspects of the subject, like me being the subject of a certain experience. There is no such gap between seeming and being the subject of my experiences and the same goes for 'occupying a certain perspective'[25] of a subject (Duncan, 2015: 2546). In both cases, we cannot doubt that there is something phenomenally present in awareness (of experience and of the subject). Therefore, I cannot doubt that I have some experiences and that it is me who is the subject of these experiences. If this seems to be the case, according to the Doubt test, we can, after all, be acquainted with our experiences and ourselves.[26]

In the following section, I will discuss what we can learn about the nature of the subject of experience from pre-reflective self-awareness (me-ness).

[25] Talk of "perspectives" can also be misleading. See Nida-Rümelin, 2017: Section 10.

[26] Russell speculates about the possibility of self-acquaintance though he is cautious since he considers it a difficult question. He admits that it is *probable* for the acquaintance of subjects to be possible, "though not certain" (Russell, 1912: 50–51). Russell says that there is acquaintance with two things in relation (subject and its sense-datum), if one is acquainted with his acquaintance with a sense-datum: "Self-acquainted-with-sense-datum". He contends that to know the truth of being acquainted with a sense-datum, we need to be acquainted with the "I", the subject. We can see a striking likeness between what Russell says about self-acquaintance and Guillot's formulation of mineness, awareness of the fact of ownership (that a subject has an experience). Some argue that we are not directly aware of ourselves but indirectly, through being aware of experiential states (Chisholm, 1969). Chisholm argued that to be "acquainted with the self as it is" just is to be "acquainted with the self as it manifests itself as having qualities" (Chisholm, 1969: 21). In support of the opposite claim, take into account what Horgan and Nichols write about *the zero point*: "The self that is present in consciousness directly and without the mediation of a self-representation — the me that is experientially present via the for-me-ness of consciousness — is directly present in experience" (Horgan & Nichols, 2016: 148). They use slightly different terminology, that is, instead of pre-reflective self-awareness, they use "non-representational self-presence" or just "phenomenal subjectivity".

2.3 INDIVIDUAL NATURE

When one has glimpsed into the essence of the subject and thus having been acquainted with oneself, one is aware that it is a thing that has experiences, according to what was said so far. One such philosopher, who argues for the kind of revelation of the subject in self-awareness is Nida-Rümelin (Nida-Rümelin, 2017: 75): "But even before such conceptualization, we are aware of ourselves as 'uniting' simultaneous and subsequent experiences. And if uniting simultaneous and subsequent experiences partially characterizes our own nature as experiencing beings, then this means that we are, in pre-reflective self-awareness, aware of ourselves as belonging to that particular ontological category; we are thus aware — in pre-reflective self-awareness — of ourselves as subjects in the following substantial sense: our nature is present to us in such self-awareness in a phenomenologically manifest way."

What this means is that through pre-reflective self-awareness, we can be aware of ourselves as entities that somehow unite our experiences and are their bearers, the owners of those experiences. Then, if this is our nature as subjects (or some partial aspect of our nature), we are aware of this aspect or characterization of our nature; we are aware of ourselves as unifiers of experiences. This is the so-called "general concept" we have of an experiencing subject, and it is based on pre-reflective self-awareness, as argued by Nida-Rümelin. The self-awareness-based conceptualization of the fact that "simultaneous instantiations of experiential properties are instantiated by one and the same subject" (Nida-Rümelin, 2017: 76) is, thus, *nature-revealing*. And what this conceptualization reveals to us could, in theoretical terms be classified as *the simple view of the self*.[27] This claim could be put in the

[27] The simple view states that simultaneous experiential properties are instantiated in one subject. Nida-Rümelin (Nida-Rümelin, 2017: Section 14) goes on to argue that pre-reflective self-awareness also gives us an understanding of our own diachronic unity, of what it means to have experiences at different moments belonging to the same subject. With it we get the *simple view about diachronic unity* and the *simple view about transtemporal identity* of subjects (Nida-Rümelin, 2012). The simple view or *non-reductive view* with respect to personal identity and diachronic unity was also advocated by E. J. Lowe.

following way: to be aware of oneself as the one who stays the same in changing experiences. It is the same subject who has all the simultaneous and past experiences and is engaged in actions.

In the self-awareness-based understanding of synchronic unity of the self, self-awareness pre-reflectively gives us the nature of ourselves as subjects, introspective knowledge that we are the unifiers of experiences. If we could conceive of a reverse case: that there is a causal connection between experiences (the *co-consciousness* relation that Dainton has posited) that makes experiences simultaneously mine or yours, such connection could not be grasped. That is, if the situation is due to the causal facts, but the subject does not conceptualize them, then the concept of synchronic unity is *opaque* in Goff's terms (Goff, 2011), and this does not seem to be right. If this is the case, then my self-awareness based understanding of my synchronic unity does not reveal to me what it is for me to have simultaneous experiences, and Nida-Rümelin rightly warns that this is an unacceptable scenario. The self-awareness based understanding of synchronic unity is nature revealing and involves self-acquaintance (and self-revelation).

It is in pre-reflective self-awareness, that we are aware of ourselves as the one who unites the experiences; this is part of our nature revealed; we are "aware of being the one single individual who has those properties at once" (Nida-Rümelin, 2017: Section 13), of being the individual with simultaneous experiences.

A related notion is that of *the phenomenal concept* of the subject/self. How should we make sense of phenomenal concepts? How can there be any concepts of the subject in pre-reflective self-awareness? Nida-Rümelin (Nida-Rümelin, 2017) tries to

---

He writes: "Moreover, the self's substantial simplicity is in no way incompatible with its manifest psychological complexity, though that simplicity does help to explain its psychological *unity.* The simplicity of the self is seen to imply that its diachronic identity — its persistence through time — is irreducible and ungrounded, and hence criterionless" (Lowe, 1996: 10). Zahavi discussed the issue of diachronic unity in his *experiential self* account, and concluded that such self has temporal extension even before obtaining narrative capacities and that "our pre-reflective self-consciousness includes some awareness of diachronicity" (Zahavi, 2014: 77).

account for this with the "general concept" of the experiencing subject. Although friends of the *Acquaintance* or the *Revelation* thesis gladly defend phenomenal concepts of experiences (experiential properties or phenomenal qualities), the same is not easily said of phenomenal concepts of subjects. There is very little literature on the topic today and substantial work is to be done in order to defend the plausibility of such phenomenal concepts.[28]

Let us now ask the important question: in order for this phenomenological argument to work, should a subject be aware of the fact of ownership or is the pre-reflective self-awareness enough? If all traits of substantival nature are revealed in self-acquaintance, then it can be inferred, very straightforwardly, that the self is revealed to be a substance. Perhaps, the property of *mineness,* in Guillot's terms, that is, an awareness between a subject and the fact that it owns the experience (where ownership is revealed), would be the most persuasive phenomenological evidence. Still, it could turn out that this property is not essential for the subject and could be absent in pathological cases.

Given the definition of a substance, three conditions need to be met — the subject/substance has to be the *bearer, unifier and individuator of experiences*. Regarding the third metaphysical requirement for something being a substance, one might find it hard to understand what *individuation* "looks" like in daily human phenomenology, if it is phenomenally present to us at all. If only the first two traits are revealed — that the subject is the bearer and unifier of experiences and not their individuator — then we need a further argument.[29] That said, I will entertain the possibility that, if the most plausible accounts of pre-reflective self-awareness and mineness are taken, some substance-like traits are revealed. So, I will present a possible argument that a substantivalist might give from phenomenology to defend its claim.

[28] Guillot argues for one "phenomenal model" of the concept of the self (I-concept), which is grounded in *cognitive phenomenology*, specifically in *the phenomenology of intellection* (e.g., Guillot, 2016).

[29] If we have a *transparent* phenomenal concept of the subject (in Goff's terminology), it is such that the whole nature of its referent is revealed. If only a part of nature is revealed, we would have a *translucent* concept.

In pre-reflective self-awareness, the experiencing individual is aware of itself as an *individual*, that is, aware of its own *individual nature*,[30] and this nature is different from the one revealed in acquaintance with the experience. Recall what was said in Section 4. Suppose one finds Goff's thesis of *Phenomenal Insight* a plausible claim, and one is acquainted with one's self as a subject and with one's experiences. In that case, one would know that the subject is something essentially different from experience (self-awareness presents in that case an essentially different content from the content in awareness of experience), and one could not confuse these two. We could then use *Phenomenal Insight* to give support to the present argument. If there is an acquaintance and if the essence of the subject and the experiences is revealed, the subject should be able, through acquaintance, to see the distinction between the subject and the experience.

If the self functions as a bearer of properties, it implies that the knowledge acquired through self-acquaintance would differ from what would be known if the self were considered a property, or an aspect in itself. This distinction in facts becomes

[30] We could explain the specific content of pre-reflective self-awareness with a reference to a *haecceity* ("thisness" or "individual essence") at the heart of the conscious individual. Could there be something like a haecceity of the subject of experiences? One version of the view was held by Swinburne (Swinburne, 1995). His position is that only *conscious* beings have haecceities and can *grasp* those haecceities. 'The property of being me, if it exists, might indeed be called a 'perspectival' property- a property which something has in virtue of being thought of or grasped from a particular 'point of view' (its own)' (Lowe, 2003: 88). Rosenkrantz (Rosenkrantz, 1993) defended the plausibility of haecceities in every object and argued that a person can grasp their own haecceity, that each individual is acquainted with himself, though haecceities of physical objects are ungraspable. Following the same intuition, Nida-Rümelin has defended that conscious individuals have a *non-descriptive* individual nature (Nida-Rümelin, 2012). One does not need to understand essences as properties. If there is a nature or essence of pain, it is not a further property that the property of pain has (Goff, 2015: 126). Positing haecceities has intuitive appeal in the case of conscious individuals (subjects). Although a proponent of the *no-self* approach, Metzinger writes about a 'distinct phenomenology of singularity, a non-sensory phenomenology of 'thisness'- for example, in the phenomenology of meditation, but also in bodily self-consciousness. If we look closely enough, we can discover the phenomenology of primitive 'thisness' in our own subjective experience. It is particularly distinct in certain non-conceptual layers of self-awareness' (Metzinger, 2011: 282). Turausky argues for this in subjects: "a non-qualitative, non-duplicable properties that uniquely individuate objects (and, in this case, subjects)" (Turausky, 2014: 249). See Lowe 2003 for a discussion on individuation.

evident in the perspectives of bundlists and substantivalists, particularly in contrast between pre-reflective self-awareness (awareness of oneself) and the awareness of experiential content (the "objects" phenomenally presented to the subject). It is crucial to recall the earlier discussion on the unique nature of pre-reflective self-awareness, characterized by its non-objectual nature. As highlighted in Section 3, the awareness of oneself and experiential content have distinct contents. This observation provides an additional argument supporting the contention that the self is a substance.

I posited that the phenomenology of pre-reflective self-awareness lends credence to the metaphysical assertion that subjects of experience are substances. To substantiate this, I drew on phenomenological evidence indicating that selves have experiences as instantiated experiential properties, with the self serving as the bearer of these properties. It's crucial to note that the Acquaintance required in my argument is exceedingly minimal—what's essential is merely the presence of the subject in awareness, known as the subject of experiences. There's no necessity for detailed knowledge of it as a substance; our awareness is limited to acknowledging the existence of a subject. Consequently, my claim is confined to asserting that we are aware of our experiences and the subject of those experiences, providing support for the substance theory. The aim was to demonstrate that the phenomenological context of pre-reflective self-awareness leans toward the substance view of selfhood without delving into the specifics of what kind of substance the self is, or which precise theory of substances should be embraced.[31]

[31] Detailed discussion can be found in Nešić, 2023b.

3. PHENOMENOLOGICAL SELVES

> I found to my horror that at times I was less conscious of myself, of my own existence, than used to be the case. This sensation was so novel that at first it quite bewildered me. I felt like asking someone constantly if I were really George Dedlow or not; but, well aware of how absurd I should seem after such a question, I refrained from speaking of my case, and strove more keenly to analyze my feelings. At times the conviction of my want of being myself was overwhelming and most painful. It was, as well as I can describe it, a deficiency in the egoistic sentiment of individuality.
>
> *The Case of George Dedlow* (Silas Weir Mitchell, 1866)

3.1 MINIMAL SELF

What is it that we cannot doubt and that we are acquainted with when it comes to ourselves as subjects of experience? We cannot doubt, at least, that we have some experiences (experiential properties instantiated) and that it is us (me, you) who have those experiences, although we can be wrong about what exactly they or some of their aspects are. One could deny that in self-acquaintance, we are presented with an individual essence, that in self-awareness, it is disclosed to one that one is a specific, individual subject that bears the mark of uniqueness. At least, this is what is claimed by some modern analytic philosophers and much of the Phenomenological tradition. This chapter deals with positions that are Phenomenological in nature and in which there is some kind of minimal self(-awareness).

Consider, for example, Zahavi's (Zahavi, 2005, 2014) view of the self. Like Strawson's and unlike some of the other views previously discussed, Zahavi's experiential self is a "thin subject" account of the self, though it is not reductive. This for-me-ness of

experience makes a difference to the subject's phenomenology. My first-person perspective is a phenomenological fact, even the pre-reflective first-person givenness of experience. Though we can imagine qualitative Perfect Twins, there is a further fact that is not entailed by those qualities, which makes them distinct: their respective individuate first-person perspectives (Zahavi, 2014). That I have these experiences does not in any way entail that I should have this first-person perspective. Deflationism in panpsychism is due to the lack of a clear notion of subjectivity. However, a "thin subject" theory of the self can still include subjectivity. Turausky (Turausky, 2014) notes that Zahavi does not posit subjects as such, just subjectivity as first-person givenness, but I think it is safe to assume with Strawson (Strawson, 2009) that when something has subjectivity, it is a subject. Subjectivity entails a subject, thinks Strawson (Strawson, 2009: 274).

> With his "minimal (experiential) self" theory, Zahavi tries to defend a third, middle-way position between substance and bundle views of the self (Zahavi, 2014: 18): "The phenomenological proposal can be seen as occupying a middle position between two opposing views. According to the first view, the self is some kind of unchanging soul-substance that is distinct an ontologically independent from the worldly objects and conscious episodes it is directed at and to which it is subject. According to the second view, there is nothing to the consciousness apart from a manifold or bundle of changing experiences. There are experiences and perceptions, but no experiencer or perceiver. A third option is available, however, the moment one realizes that an understanding of what it means to be a self calls for an examination of the structure of experience, and vice versa."[32]

[32] The experiential self that Zahavi defends has temporal extension and is something that can be shared by many experiences, although there may be interruptions of the stream of consciousness (an unconscious episode of sleep and coma). These characteristics set Zahavi's view apart from the bundle view, though he still does not posit an additional self to account for the diachronic unity and personal identity. That said, some kind of *awareness of diachronicity* in pre-reflective self-awareness is preserved (Zahavi, 2014: 77). This is explicitly stated by Zahavi: "Whether the same experiential self is present in two temporally distinct experiences depends on whether the two experiences in question partake in the same dimension of mineness or for-me-ness" (Zahavi, 2014: 72).

The self is seen as a feature or function of the givenness of experience, as a dimension of experience that defies both elimination (in a bundle) and inflation to a substance. This could be called the "*thin subject"* view.

Following what has been said in the discussion of Guillot (Guillot, 2017), it has become clear that to have a phenomenological and a metaphysical claim about the subject of experience, one needs something more than just *for-me-ness* of experience. Self-awareness is thus needed for this (me-ness). So, to have any introspective knowledge about the subject, we first need to have self-awareness. There is, perhaps, no mineness as a feature of the experience, but there is "mineness" as awareness between a subject and the fact of ownership. If there were just something like *for-me-ness* in experience (which is what Zahavi usually assumes to be *mineness*), this would be insufficient to support the subject as substance claims.

Guillot criticizes Zahavi's position by showing that he moves from an epistemic to a phenomenal and metaphysical thesis, "from the 'self-manifestation' of experience (*for-me-ness*) to a phenomenal access to the self (me-ness)" (Guillot, 2017: 50). Zahavi makes an illegitimate move based on an unjustified assumption of equivalence because he conflates *for-me-ness* with *me-ness* and ends up claiming that a property (for-me-ness as a quality of experience) *is* the "minimal self*"*.

Zahavi moves from an epistemic to a phenomenal and metaphysical thesis, "from the "self-manifestation" of experience (for-me-ness) to phenomenal access to the self (me-ness)" (Guillot, 2016: 50). He makes this leap because he conflates for-me-ness with me-ness. The problem with Zahavi's "thin" or minimal self account seems to be that it puts the self and experience too close without making the necessary phenomenological and metaphysical distinctions. If it is not to be judged by metaphysical reasons that the experiencer and experiences are not identical, then this can be inferred from phenomenological datums of self-awareness and content of experience-awareness. Zahavi's for-me-ness seems to encompass several different notions, and this problematic for-me-ness leads him to conclude that there is a minimal self. Not making a clear distinction between for-me-ness and me-ness

(or mineness of Guillot, 2016) in phenomenology gives philosophical accounts of mineness (from Zahavi to Guillot), pre-reflective self-consciousness (Nida-Rümelin) and acquaintance (Gertler, Goff). I think it can be shown that if one is acquainted with oneself, i.e. if one has self-acquaintance and acquaintance with one's experiences (so one has self-awareness and awareness of experiences), one acquires introspective knowledge that oneself is a substance. To do this, one would have to demonstrate that if the subject is self-acquainted, then this revelation of its nature in self-awareness gives it justified introspective (phenomenal) knowledge that it is an experiential subject which has experiences (experiential properties) and so is of the substantival kind.

Self-consciousness and subjectivity that unify experiences are the pre-reflective first-personal givenness of experience, mineness or ipseity (Zahavi, 2005). For Zahavi, this is the experiential ("minimal") self. Self is not something detachable from its experiences. This is the middle way of "the phenomenological proposal", a view posited between regarding the self as an entity distinct, i.e. separated from experience, and a view that the self is a manifold or a bundle of experiences; neither can it be detached nor reduced to experiences (Zahavi, 2014: 18). And as Zahavi explains in his discussion on act-transcendent identity of the self in Husserl, the self cannot be given as identical in just one act, it is known as identical to itself in the synthesis of the manifold of experiences that come and go (Zahavi, 2005: 131). Thus, self/ego is the abiding dimension of first-personal experiencing, as Zahavi sometimes formulates it.

Fasching views the first-personal givenness as a dimension, and as such not the result of relations between experiences, "but […] what makes them possible": 'one awareness' (the togetherness of the manifold synchronically co-conscious experiential contents) is not a result of any relations between the experiential contents, of some synthesis of them… the character of a dimension in which the contents, with all their relations, have their presence in the first place (Fasching, 2009: 143–144). This would be to conceive of a self as a phenomenal space. However, even if the self is imagined as a kind of space that holds the experiences, it would have a substantivalist interpretation, or so I would argue.

Dainton has considered and rejected the notion of a "subjective space" (Dainton, 2008: 141–145). He equates it with an A-thesis or pure awareness thesis. He denies that selves are identical with phenomenal spaces, as proposed by Stephen Priest, because such a phenomenal space would have to be substantival rather than relational, and this cannot be the case. Dainton thinks that for it to be substantival it has to "have some introspectively discernible qualitative phenomenal features of a recognisably spatial kind", and he argues that it does not have such a phenomenal feature, i.e. that there is no mineness (Dainton, 2008: 101–145).

Authors like Zahavi and Fasching seem to argue against an account of a subject as a substance that is oversimplified and not the only one possible position that one can assume towards the nature of the subject. As Zahavi has himself noted, the no-no-self view also comes in various flavours and strengths. Zahavi's notion of experiential self is too thin and deflationary because he puts too much emphasis on the first-personal character, and this hides the subject or ego as a "mental thing", a something, and not a way a thing is, ego as an individual substance and not a mode of a substance or a mode without a substance. In the end, I think some of these authors are arguing against the "bare particular" view of the subject (but also against the bundle view, such as Parfit's).

Some argue that it is a category mistake to claim, as Zahavi does, that mineness, as a feature or property of the experience, is the experiential self (pointed out by Siewert). On the other hand, we can safely claim that mineness implies an experiential self. Mineness, as Zahavi understands it, could indicate that there is something more than experiences and their relations to a subject of experience. Guillot (Guillot, 2016) proposes that subjective character refers to several distinct notions that are being confused by some authors: for-me-ness (a relation of awareness between a subject and an experience), me-ness (a reflexive relation of awareness a subject has to itself) and mineness (a relation of awareness between subject and a fact that it owns the experience), and all these are about relations of awareness between a subject and its experiences. What Zahavi seems to have in mind when he talks about mineness is for-me-ness.

A bundlist would deny that there is self-awareness or mineness in any of the forms proposed earlier. What one needs to do is to anticipate the would-be bundlist response to substantivalist arguments: they could say that the *core bundle* plays the role of the subject (essential properties of the bundle) and that this fact is indiscernible from substantivalist phenomenology, that is, it feels the same phenomenologically as being a substance. The bundlist appears to have a more economic theory of the self because it posits only one category.

One such bundle theory of the self is Barry Dainton's *phenomenal self* theory.[33] It is claimed that by postulating something like a phenomenal background, a bundlist could explain self-awareness. Still, I would argue that, on the grounds of acquaintance in awareness, the difference in metaphysics produces a difference in phenomenology.

Dainton understands mineness as a *meish* quality to experience (Dainton, 2008: Ch. 8) and asks if it exists in phenomenology. As it was argued in earlier sections, such an understanding of mineness and self-awareness is ill-conceived and misleading. The phenomenal background has an inner component consisting of "bodily experience, thoughts, memories, imaginings, and emotions", that is, of *experiences*. The inner phenomenal background creates (constitutes) the feeling of being me or you, the ambient "sense of self". The natural intimacy of "mineness" is gained when a new experience is incorporated into this background. Slors and Jongepier (Slors & Jongepier, 2014) argue that the mineness of experience is a product of the *external* structure of experiences in their *reductionist coherentist* account. It would appear that what they consider mineness is very different from what Zahavi has in mind, or what we find in Guillot (Guillot, 2017). Although Slors and Jongepier agree with Zahavi that thoughts have first-personal givenness, the mineness they defend has nothing to do with

[33] Donnchadh O'Conaill writes how Dainton "has developed a sophisticated version of the bundle theory", distinct from the classic bundle theory of Hume, one in which the subject is a bundle of *capacities for experiences* and not a bundle of experiences themselves (O'Conaill, 2020: 1–2). O'Conaill argues in his paper that Dainton's co-consciousness, as a relation of experiential togetherness, presupposes a common subject of experiences and that the identity-conditions of experiential capacities cannot be specified without their subjects.

how it is usually conceived. It is the sense of familiarity or coherence of a certain experience with other background experiences. It is similar to Guillot's third property of mineness. Also, there is no real phenomenal datum to it, and Slor and Jongepier are very explicit about this, saying that there is "an absence of a further experiential feature" (Slors & Jongepier, 2014: 194).

That the experience is mine is explained by the *co-consciousness* of this experience with the inner component of the phenomenal background. This background is the phenomenal *I* present in consciousness (Dainton, 2008: 243). He points out that Parfit has also advocated "the reductionist view of our *sense of self*".[34] Dainton thinks that there is no special awareness of the self as a thing, awareness of the subject as a subject, as that which is experiencing something, *the experiencer*. Apart from this being metaphysically problematic, it seems to me that it is phenomenologically unjustified. Usually, Dainton's C-system and the *phenomenal self* theory are attacked by metaphysics. I think Dainton's theory could be criticized from a different (and arguably more plausible) understanding of mineness as self-awareness, as a real *sense of self*, that is, as pre-reflective self-awareness of the subject as a subject of experiences.

3.2 DIMENSION OF SUBJECTIVITY

According to Gurwitch's classic distinction[35] (Gurwitch, 1941), theories of consciousness can be categorized as either egological or non-egological. In an egological theory, the proposition is that when observing a film, there exists a self or ego that is conscious of itself as an ego watching the film. It asserts

[34] Billon employs depersonalization cases to challenge Dainton's inner background theory of mineness. Depersonalisation can affect *all* conscious states, even those in the background. This suggests to Billon that mineness is explanatory prior to co-consciousness. Billon contends that when we have unimpaired basic self-awareness it still cannot inform us on the nature of the self (Billon, 2017: 6).

[35] "The ego is to psychic objects as the universe is to material things: the former and the latter have to be taken as infinite synthetic totalities. By the mere fact of their coexistence the psychic objects group into an organized unity; this is the ego." (Gurwitsch, 1941: 336).

that every experience inherently involves a subject. Conversely, a non-egological theory perceives the experience as an anonymous awareness within consciousness, lacking a distinct ego observing the film.

The Heidelberg School, Henrich and Pothast, and Brentano and Sartre defend non-egological theories of consciousness and self-consciousness. Self-consciousness is anonymous according to them. Pre-reflective self-awareness is originally egoless as these authors claim (Henrich, 1970; Pothast, 1971; Zahavi, 2005; Miguens, Preyer & Morando, 2016). It is an anonymous dimension. Henrich proposes a fundamental separation between self-consciousness (which he terms *Bewusstsein*) and self-knowledge (which he terms *Selbstsein*). Self-consciousness should be considered an original dimension of openness or an anonymous "field"; there is no I inhabiting it, no conceptual goings-on taking place there. The current "Philosophy of Mind" would speak of self-awareness. Self-knowledge (Selbstsein) has — as Kant's "I think"– an I as centre; from there conceptual operations arise" (Henrich, 1971/2007: 12, 17; cited in Frank, 2016: 40).

Pothast's argument is as follows: if ego is something that has experiences to be opposed to experiences or something that is distinct from experience, then why is the ego's awareness of experiences an instance of self-awareness (Pothast, 1971; Zahavi, 2007: 277)? Even if they are not identical, why should this be the case? Subject "knows" itself by having experiences given to itself (givenness); it knows itself implicitly, pre-reflectively, without turning its eye to itself.

"Whatever this self might be, it is at least an active principle of organization in the field of consciousness" (Henrich, 1970: 20). "But the knowledge the self has of itself is a fundamental situation which can be explained only by its function as an organizing principle of an egoless consciousness." (Henrich, 1970: 21). Later on, Henrich describes consciousness as an occurrence of a special kind that makes other occurrences possible, and that is why it is a dimension or a medium, it is "actually only the grounds of the possibility of a type of relation, and as such, one might better describe it as a dimension or medium: it cannot exist without a system of relations between elements given to consciousness

which are themselves different from consciousness." (Henrich, 1970: 21–22). He also states that the list of its characteristics should also include that it is an exclusive dimension. How can it be exclusive if it is anonymous is not clear. It seems Zahavi's self is too minimal to be an egological self.

Following Zahavi's interpretation of Husserl, it can be concluded that there is no total anonymity. Self lives in anonymity, for Husserl, not because there is no self and it is an anonymous dimension (The Heidelberg school), but because it is not thematically conscious of itself. There is pre-reflective consciousness, however. "And when I do thematize myself in a reflection, the very act of thematization remains unthematic." There is always some "part" (not really a proper part, some aspect) of the self that stays anonymous, unilluminated, forever sombre and dark. In this interpretation (Husserl's), there is pre-reflective individuation, though there is anonymity due to a lack of thematic reflection. This would be the stronger claim that there is no self-awareness and individuation (Zahavi identifies Merleau-Ponty). I would say that Henrich et al. also fit into this category; they have this claim. With them, there is no pre-reflective ego or self. Self-awareness (pre-reflective) is egoless and anonymous.

Now, just because the subject is anonymous, this does not mean there is no ego in it. It is just ego given to itself in an anonymous mode (way), non-thematic. So there could be "anonymous" individuation without the thematic individuation of reflection, so its pre-reflective does not find itself as an object. Zahavi criticizes Henrich, Pothast, Sartre, Hume, Nietzsche, Metzinger and others who proposed there was no self that they assumed this based on criticism of a specific notion of a self, that of a distinct entity, substance. Nevertheless, there can be other, different notions of the self. He distances himself from the substance view and the no-self view of Henrich, saying that its notion of self as activity is inadequate. However, Zahavi's notion is very close to Henrich's and almost as "thin" as no-self. It is not a substance and not a bundle of experiences. Even if we assume that the self is nothing but a dimension (Zahavi, 2005; Klawonn, 1987; Henrich, 1970; Fasching, 2009, 2011), it is still not reducible to content (experience) and its relations; it makes them possible in the first place.

3.3 DOUBLE-ANONYMOUS SUBJECT

Merleau-Ponty's understanding of subjectivity brings an exciting and different angle, but his views are rarely discussed in the literature, or at least not enough. Positioning his approach within other phenomenological theories and modern variants of those views would be quite beneficial for our present understanding of selfhood, especially for enactivist theories of the self, as we will see later.

All the phenomenologists are approaching the subject and subjectivity from what today is called an embodied perspective, meaning that there is no consciousness and awareness which is not mediated through the body, the so-called "lived body", which can shift between being a living subject that is engaging with the world (*Leib*) and a material object (*Körper*). Merleau-Ponty has emphasized the concept of *anonymity*, the "anonymous life of the body" in *Phenomenology of Perception* (Merleau-Ponty, 1945/2002), thus, according to some commentators, breaking away from Husserlian phenomenology (abandoning transcendental egology). This concept of anonymity has been very controversial. Contrary to some readings, e.g. Barbaras' (Barbaras, 1991), Heinamaa (Heinamaa, 2015) puts forward an apologetic interpretation of Merleau-Ponty and his view on the anonymity of the subject, defending the view that "by anonymous perception, Merleau-Ponty means neither self-less experience nor any experience with a general or collective subject fusing together personal selves" (Heinamaa, 2015:125).

According to her, Merleau-Ponty's anonymity does not deny self-awareness (cogito), and this anonymous subject is singular, not plural or general, and in each perception, there is a personal and anonymous element (Merleau-Ponty, 2015: 126). He says that through reflection, we find "along with the perceiving subject, a pre-personal subject (…) [*un sujet prepersonnel*]." Merleau-Ponty (Merleau-Ponty, 1945/1993: 405; 1945/2002). So, there are two of them, yet he states that the subject is "someone", a singular, and there is a "couple or a pair, me and an unspecified, nameless or un-named singular, someone else with me" (Heinamaa, 2015: 129). The one who "perceives with me" is

the sensing-moving living body. Furthermore, sense organs of the body form a dynamic and integrated system, so this self-awareness is not a simple, complex whole, "I am given to myself as a sensing body, and my sensing body incorporates an integrated system of my sense organs and sensory functions." (Heinamaa, 2015: 131). Merleau-Ponty calls sensations "little subjectivities" (Merleau-Ponty, 1945/1993: 279), but they are not determined by me, but by the sensible, perceivable things, and that is why there is anonymity in the movements and perceptions of my body.

For Merleau-Ponty, our subjectivity has "double anonymity"; the absolute subject of experience is "anonymous", but we are also the modifications of the styles of perceiving that we inherit from our ancestors, both human and animal (Heinamaa, 2015: 135), and such a subject cannot be even called the self or ego. In Merleau-Ponty, there is then a perceptual coupling[36], "Perceptual life is an ongoing active process of "movement" that engages a "field" and thereby brings stable unities (objects) into relief: "I perceive a thing because I have a field of existence and because each phenomenon that appears polarizes my entire body, as a system of perceptual powers, toward it" (Merleau-Ponty, 1945/2002: 332) and the same goes for dynamic coupling of intersubjectivity/intercorporeality through which the We is formed (Walsh, 2019).

[36] The same is the case in enactive theories of perception.

4. SELVES ARE EVERYWHERE!

> Suddenly every object in my field of vision took on a curious and intense kind of existence of its own; that is, everything appeared to have an "inside"– to exist as I existed, having inwardness, a kind of individual life, and every object, seen under this aspect, appeared exceedingly beautiful. There was a cat out there, with its head lifted, effortlessly watching a wasp that moved without moving just above its head. Everything was *urgent* with life... which was the same in the cat, the wasp, the broken bottles, and merely manifested itself differently in these individuals (which did not therefore cease to be individuals however). All things seemed to glow with a light that came from within them.
>
> *Mysticism and Philosophy* (Stace, 1961: 72)

4.1 IT'S ALIVE! (PANPSYCHISM)

There has been a revival of interest in both the plausibility of self-consciousness and panpsychism in contemporary philosophy of mind. Nevertheless, the importance of subjectivity or pre-reflective self-consciousness in experience has been neglected in panpsychist accounts of consciousness. I will, thus, go on to argue that we should look to phenomenology to understand better and be able to solve the proposed combination problems that one can encounter in panpsychism. As an extension of materialism, panpsychism holds that physical matter does not generate consciousness but is already endowed with it. The view harbours some serious core problems that are now commonly referred to as the *combination* problems: how consciousness from lower levels generates our level consciousness, how subjects sum to yield a new subject or how experience makes a unified subject of experience.

To put it roughly, panpsychism is a metaphysical theory that claims everything in nature has a modicum of consciousness. Though it may seem counter-intuitive or even downright crazy, panpsychism was conceived in order to overcome the deficiencies and problems of both physicalism and dualism.[37] Panpsychism should successfully answer the hard problem.

Chalmers has made the canonical distinction between easy and hard problems of consciousness (Chalmers, 1995). The really tough problem that needs answering is why and how we become conscious in a materialistic world. How does it come about from physicalist systems? "How does the brain process environmental stimulation? How does it integrate information? How do we produce reports on internal states? These are important questions, but to answer them is not to solve the hard problem: Why is all this processing accompanied by an experienced inner life?" (Chalmers, 1996: 12). The so-called easy problems of consciousness (like the focus of attention, the deliberate control of behaviour, the difference between wakefulness and sleep) are problems with the performance of *functions*. The hard problem is not about functions. "But how and why do physical processes give rise to experience? Why do not these processes take place 'in the dark' without any accompanying states of experience? This is the central mystery of consciousness" (Chalmers, 2003: 103–4). "How can technicolour phenomenology arise from soggy grey matter?" as famously exclaimed by Colin McGinn (McGinn, 1989: 349).[38] The hard problem comes from human phenomenology — it is like feeling the warm touch of the partner's hand on your face (or having the phenomenality of bodily sensation, feeling and thought). So, it is posited that consciousness is characterized by a special kind of *phenomenal properties* (or *qualia;* Chalmers, 2003). Theories of consciousness should give an answer to the hard problem of consciousness. The question to ask now is whether IIT, as a neuroscientific account of consciousness, answers the hard problem.

[37] For more on this see Chalmers, 2015.

[38] Cf. Strawson's (Strawson, 2008) "the hard part of the mind-body problem" and Levine's (Levine, 1983) "explanatory gap".

There are several forms of panpsychism debated today. Constitutive panpsychism, where macroexperience is (wholly or partially) grounded in microexperience and emergent panpsychism, in which macroexperience is emergent from microexperience (Roelofs, 2015, 2019). Russellian panpsychism involves the thesis that quiddities are the fundamental categorical bases of relational physical properties. Panprotopsychism implies the thesis that fundamental physical entities are of the proto-conscious kind (Chalmers, 2015: 252–259). Mørch (Mørch, 2014) presents emergentist version of panpsychism. In addition, there is the *cosmopsychist* view (cosmopsychism), that the universe itself instantiates experiential properties (see Bruntrup & Jaskolla, 2016). Constitutive panpsychism seems to suffer from difficult new problems, so-called combination problems (cf. Coleman, 2013; Chalmers, 2016; Goff, 2017: Chapters 7 and 8) that relate to the issues surrounding the emergence of macro-consciousness from micro-consciousness. The analogous hard problem of combination poses the question of how macro-subjects (o-subjects) come from micro-subjects (the subject summing problem). Cosmopsychism has a *decomposition/decombination* problem (how smaller subjects come about from an absolute one subject). In previous work, I tackled the thorny problem of subject-summing (Nešić, 2017) and an emergentist panpsychist interpretation of IIT (Nešić, 2018a).

If IIT is interpreted as a physicalist theory, then it would seem to be afflicted by the same problem of explaining consciousness as other physicalist and materialist positions. In contemporary literature on panpsychism, it is argued that the position comes with a set of its own, new and sturdy problems, those of *combination*. The most plausible form of panpsychism (the one that avoids emergence), the *constitutive* type of panpsychism, runs into difficulties.[39] Constitutive panpsychism posits that macro-consciousness is grounded in micro-consciousness;

[39] For more on this, see Coleman, 2013. It was Goff who argued that a set of subjects does not a priori entail the existence of another subject (Goff, 2009). Coleman goes on to strengthen Goff's claim and says that it is metaphysically impossible to combine subjects to yield new subjects, and this leads some to dismiss constitutive panpsychism as to avoid *brute* emergence.

macro-experience just has those micro-experiences as parts, and it inherits their properties. The combination problem for this metaphysical theory arises when we try to understand how o-consciousness or macro-consciousness (that we pre-theoretically know in ourselves) comes from fundamental micro-consciousness of its parts. The problem is particularly hard to (re) solve when it relates the combination of micro-subjects (even if it is intelligible that such a subject can exist!) to macro-subjects of human beings.

Russellian monism (RM) has recently been defined by Chalmers (Chalmers, 2015: 262) as the view that: "structural properties in physics do not constitute consciousness, but quiddities (perhaps along with structure) constitute consciousness. The view is Russellian because of the appeal to quiddities and their connection to mentality. It is a sort of monism because the world in this view consists of quiddities connected by laws of nature." The most plausible type seems to be Russellian identity theory in which dispositional properties are identical to categorical properties ("powerful qualities" view of properties).

4.2 CAN WE PUT OUR MINDS TOGETHER?

Although Russellian Monism can be an appealing theory addressing the mind/body problem, as I said before, it suffers from combination problems. These problems come about when we try to understand how macro-consciousness (*o-consciousness* that we pre-theoretically know) comes from fundamental micro-consciousness. The issue is challenging when it relates to combining micro- into macro-subjects, called *the subject-summing problem*. One easy way out of the subject-summing problem for the panpsychist is to deny the existence of any micro-subjects since the only type of subjectivity we phenomenologically know is one of human beings. If we postulate the existence of micro-subjects and we deny the possibility of subjects summing together to form a larger subject (combinationism) then we have an almost unsolvable problem on our hands: the very hard problem of the subject combination.

When it comes to solving the combination problem in Russellian monism, Chalmers argues that the most promising position is that of the constitutive panpsychism (or constitutive Russellian monism). According to constitutive panpsychism, macro-experience is wholly or partially grounded in or constituted by micro-experience (Chalmers, 2016: 25). Chalmers argues against Russellian panpsychism in which emergence is considered as a brute fact, where o-experience is strongly emergent on micro-experience. Goff stresses that such brute emergence is incoherent, and we should exchange brute emergence for intelligible causation. He comes to the position of intelligible emergentism and combines it with Russellian panpsychism (Goff, 2015: 394). Goff argues that if there is a solution to the combination problem, and there is a priori entailment from the micro-phenomenal to the o-phenomenal facts, then intelligible emergent Russellian panpsychism is a good account for o-consciousness.

Coleman (Coleman, 2013) has pointed out that there is an "internal tension" in panpsychism, and he argues against the possibility of subject-summing. Constitutive panpsychism was driven by an aversion to emergentism[40], but in the end, it seems that has to yield to some kind of emergence in order to account for the production of high-level subjects. He argues that if panpsychism re-sorts to emergence, then classic physicalism could prove to be more plausible. Accepting emergence could be taken as a betrayal of the original intentions of panpsychism.

Coleman is a proponent of *panqualityism*, position in which ultimates are (absolutely) [1] intrinsic *qualities*, but they have no experience; they are not *subjects*. There are two basic ontological hypotheses (that ultimates *have/are intrinsic qualities* and that they have *experience*) that come from accepting experience at the base levels of reality, *contra* orthodox physicalism. The later hypotheses is stronger and leads to panpsychism. The weaker leads to some form of neutral monism. Sometimes the term panqualityism is used for such a position in which intrinsic qualities

[40] Nagel's famous argument for panpsychism includes a Non-emergence premise: "P4. Non-emergence: All high-level properties of a composite intelligibly derive from properties of its constituents plus their arrangement" (Nagel, 1979: 181–182).

are the ontological bottom ground but it is considered they have no experience.

Coleman takes such a route thinking it a golden middle way between orthodox physicalism and full-fledged panpsychism. He argues that the subject combination problem is unsolvable in panpsychism. Panqualityism avoids the troubles of panpsychism by having ultimates not be conscious subjects, though they are intrinsic qualities that orthodox physicalism can't explain. So if the subject combination problem proves to be the hard problem for panpsychism, then Coleman's panqualityism will count as a more appealing position. But there are greater problems with panqualityism that make it unappealing. And it seems to me that panqualityism is under the same threat as emergentism, for it is difficult to understand how non-experiential qualities become experienced, and where the place for subjects is.

The subject problem in panqualityism seems to be unsolvable. Take in consideration that in panqualityism experience is functionally reducible, as if somehow these non-phenomenal qualities managed to become aware (experience) of themselves by themselves and so experience is a relational property. Since subjects are not primitive entities in panqualityism then they have to be emergent. However, panqualityists thinks that the step from non-experiential to experiential is not a giant leap. One possibility that Coleman argues for is a *higher-order thought theory of consciousnes*s (HOTT) and he advises such a relationist solution to panpsychists turned neutral monists. Qualities come to conscious awareness when they are represented by a higher-order thought. This is how *qualities* become *phenomenal* properties. However it is hard to see how it deals with phenomenology in a satisfactory manner and how exactly awareness and subjectivity come into play. Something is missing here. Coleman considers the case of Humphreys' fusion (and combinatorial fusion of Seager) and concludes that it can't help the panpsychist, because there is an emergence of macro-subjects in such an account. "If points of view are annihilated so as to produce a macro-subject, the macro-subjectivity is not a structural property" (Coleman, 2013: 19). So the micro-subjects cause the macro-subject to come into being, and this is what happens in emergence.

Coleman claims this happens in combinatorial fusion and is a form of emergence.

The problem with emergent forms of Russellian panpsychism is that they breach micro-physical causal closure (Goff, 2015: 396), though there is no such problem with macro-physical causal closure. For the emergent RM, macro-physical is nothing over and above macro-phenomenal, so there is no problem with macro-physical causal closure. But since o-experience is over and above everything micro, and micro-closure is accepted, then there is a problem with the causal efficacy of macro-experience (o-experience). An emergent account of RM that would accommodate micro-physical closure would have an obvious advantage.

The amalgamation of individual subjects appears to pose an insurmountable challenge for constitutive panpsychism, and, moreover, it seems fundamentally unsolvable. Goff (Goff, 2009) contends that the mere existence of one set of subjects does not, a priori, imply the existence of another subject, and the mechanism through which this might happen remains elusive, although there is a possibility it could occur in some yet unknown manner. Coleman, while acknowledging Goff's argument, takes it a step further by asserting that combining subjects to generate new subjects is metaphysically impossible. According to Coleman, such a combination would entail a form of brute emergence, a concept panpsychists seek to avoid, given their initial intent to steer clear of emergence. Consequently, Coleman dismisses the viability of constitutive panpsychism and suggests that panpsychists should instead adopt a neutral monist stance. Although I find Coleman's stronger version of the no-summing argument quite compelling, I diverge from his conclusions and proposed solution. While I acknowledge that the argument challenges constitutive panpsychism, I believe it doesn't undermine other forms of emergent panpsychism, such as the one presented here.

If one were to embrace fusion, as advocated by Humphreys (Humphreys, 1997) and Seager (Seager, 2010), to derive macro-subjects from micro-subjects, it would entail a form of strong, radical emergence. Both Seager and Mørch (Seager,

2010; Mørch, 2014) incorporate some version of such emergence into their theories. In the context of fusion, subjects at a lower level generate a new subject while ceasing to exist themselves—they undergo annihilation in the process. Mørch puts forth an interpretation wherein subjects partially survive; they undergo change due to the whole entity of which they are parts but persist in existence, especially considering the absence of fusion in the brain. In this account, there is no emergence of a truly novel macro-subject in the strong sense, separate from the micro-subjects that constitute it, even though these micro-subjects undergo some degree of change. The nature of fusion implies radical emergence exclusively, offering no support for the constitutive panpsychist. In Mørch's framework, there is no occurrence of strong emergence, as the micro-subjects survive, but simultaneously, there is no manifestation of a macro-subject in any robust sense, making this position deflationary. The combination of subjects seems to be an insurmountable problem for constitutive panpsychism. More than that, it could be insoluble in principle.

Goff (Goff, 2009) argues that a set of subjects does not a priori entail the existence of another subject as stressed before; we cannot see how this would happen, though there is a possibility open that it might happen in some, to us, presently unknown way. Coleman, while acknowledging Goff's argument, goes on to strengthen his claim and say that it is metaphysically impossible to combine together or assemble subjects to yield new unique subjects and this, as a consequence, rules out constitutive panpsychism as a contender position. If this were the case, it would imply some kind of brute emergence, and panpsychists don't want this because it is a position that they originally tried to avoid. This is what ultimately drives Coleman to deny the plausibility of constitutive panpsychism and argue for neutral monism as the best position to take. Coleman considers this to be a golden middle way between orthodox physicalism and full-fledged panpsychism and defends the view referred as panqualityism, a position in which the ultimates are qualities.

4.3 DEFLATING THE SUBJECT

Deflationist views about subjects of experience dominate the landscape of contemporary literature on panpsychism. Deflationary subjects of experience are not persistent through time as we pre-theoretically conceive of them. They are not diachronically unified, though they can have synchronic unity. Deflationary views were defended by Hume, James, and Parfit (Hume, 1739–40; James, 1890; Parfit, 1971). Modern panpsychist accounts of Roelofs, Mørch, Strawson, Coleman, and Seager (Strawson, 2009; Seager, 2010; Coleman, 2013; Mørch, 2014; Roelofs, 2015) all have deflationist traits.

Chalmers (Chalmers, 2015, 2016) shows that one possible reaction to the combination problem is to deflate the subject. *Prima facie,* it is an appealing strategy. But it seems that deflationism about some of the main concepts of panpsychism (subjects, awareness) cannot on itself be a solution to any of the combination problems, though it can help us get to those solutions (together with phenomenal bonding, for example). "Either denying that experiences must have subjects at all, or at least denying that subjects are metaphysically and conceptually simple entities" (Chalmers, 2015: 271). Though he sees it as a conceptual truth that experiences must have subjects who have them, he finds the second denial untenable. Opting for deflation seems like a natural choice in compositional panpsychism/panprotopsychism. Deflation garners special attention regarding the possible solution to the subject combination problem. If we deny the deflation of the subject, there can be no combination. Deflation can range from denying that subjects are metaphysically primitive entities (reduction) to complete eliminativism of such entities. Given deflationism, there is greater plausibility in the composition of subjects from other subjects and the composition of subjects from experience. Are we primitive subjects of experience, Edenic Subjects? Should we imbue our Subjects with sedulous subjectivity? If we are something like Edenic subjects, this would eliminate constitutive pan(proto)psychism.

Eliminativism of subjects can be found in some neutral monist views, like in the theories of Russell, Mach and James.

The eliminativist views face combination problems, and even less extreme forms of deflationism that view subjects as composite and derivative still have the subject-summing problem (Chalmers, 2016). There is also a deflation of awareness. James is eliminativist about such relation, and Coleman defends a reductive, functional analysis of awareness (Coleman, 2013).

Panqualityism is subject to the "nonsubject/subject gap" problem (Chalmers, 2015: 272) and tries to "patch it up" with deflation. In such a view, quiddities of microphysical properties are qualities. When there is awareness of qualities, they become phenomenal properties. In panqualityism, subjectivity is not essential to qualities. Some panqualityists reject subjects of experience altogether (eliminativism), while others think they are constituted by qualities in certain relations, as with Coleman's solutions and the higher-order thought theories of consciousness. Panqualityism of Coleman, in which the basic, intrinsic properties are qualities as "unexperienced qualia", harbours the conceivability of awareness zombies (Chalmers, 2016). Attempts to "functionalize" awareness eventually eradicate the phenomenology of awareness, as is the case in panqualityism. What all this comes down to is the problem of radical emergence. The hard problem of consciousness originated as a result of the unintelligibility of the radical emergence of experiential from physical (as completely non-experiential). Panpsychism came as an answer, assuming that the experiential can only come or emerge (in a non-problematic way) from the experiential (*No-Radical-Emergence Thesis*). However, the explanatory gap reappears in the panpsychism's and panprotopsychism's combination problems. There seem to be problems of radical emergence of subjects from qualities, of subjects from experience and of subjects from other subjects. Chalmers (Chalmers, 2015) claims there is an explanatory gap between qualities and awareness and a gap between qualities and experience on account of the conceivability argument. I will not be concerned with eliminativist positions, only with deflationist ones, particularly with those that reduce subjects to experience.

So, there opens an explanatory gap between subjects and experience, and some philosophers try to answer it with deflationism by reducing subjects to experience. If we argue against

deflationist reduction on the grounds that there is something like irreducible subjectivity, this could push us towards giving up constitutive panpsychism. We should be wary of reduction in these matters because we might not know well that what is being reduced, to take a cue from Nagel (Nagel, 1974). I think the deflationist reductive approach can be challenged by appealing to the first-person givenness of experience, which is what I will argue for.

Let us examine several representative deflationary views of the self/subject that are endorsed by modern panpsychists. Parfit's theory, though it is not panpsychist in nature, presents a reductionist account of personal identity and of the self (subject). Mørch (Mørch, 2014) cites Parfit's and Strawson's deflationist accounts that are useful in solving the combination problem. In the next sections, I will make apparent the shortcomings of several deflationary positions and offer a better solution. One of the reasons we think of ourselves as subjects in a strong sense is the intuition about the persistence of personal identity. Cases of personal identity breakdown are taken to support deflationary views of subjects, and Parft's thought experiments contribute to subject reductionism. I will consider one example of a panpsychist deflationist theory of the self.

Parfit (Parfit, 1971) imagines fission (division) and fusion cases. According to him, there is no personal identity, just survival. These cases show him that when it comes to survival everything is a matter of degree; survival is not a matter of all-or-nothing. Such cases involve the imaginary scenario in which someone would have voluntary control over his corpus callosum, a bridge between the hemispheres of the brain. Under control, it could be temporarily disconnected, and each hemisphere would have its own independent stream of consciousness until they are connected again. For example, each hemisphere could do a separate math calculation. This would be Parfitian psychological fusion. These cases serve to weaken the rigidity of personal identity. He just asks for psychological connectedness between "persons". Parfit's psychological criteria for personal identity are non-exact similarity and causal connectedness (Mørch, 2014).

Parfit concedes that the subject is single: "in each of my two streams of consciousness, I would believe that I was now, in my other stream, having thoughts and sensations of which, in this stream, I was unaware" (Parfit, 1984: 246–8). He is a reductionist when it comes to the self and personal identity. He prefers to call his position Constitutive Reductionism. This means he thinks that we are distinct from our brains but not as separately existing entities (Parfit, 1998: 218).

For Parfit "a person's existence just consists in the existence of a body, and the occurrence of a series of thoughts, experiences, and other mental and physical events" (Parfit, 1995: 16). So not to be just a Reductionist, he claims though a person is distinct from the body and thoughts and experiences, a person just consists of them. Although he is not a panpsychist and Reductionists do not believe in mental substances, they can still be dualists, according to Parfit. Mental or not, his persons are not distinct entities, they just consist of the stream of experiences and thoughts, whatever their nature may be. This is why we may consider him a proponent of a deflationary view of persons (subjects). Contrary to the common belief, Parfit argues for the unimportance of personal identity. He thinks that what is important is psychological continuity and psychological connectedness between different parts of a person's life (like between me now and some future me). And these relations are a matter of degree, something that is not the case with identity.

Parfit envisions experiences as wholly "impersonal". There seems to be no notion of the subjective in this kind of Reductionism and no distinct subject. What would then distinguish my experiences from someone else's? It seems to me that the real criterion for personal identity is and should be the subjective character of consciousness, the primitive self-experience, that is, the tell-tale of someone's subjectivity, of someone being a subject of experience.

4.4 SESMETS

We find Strawson's position on the question of the subject somewhere in the middle, between pro-selfers and anti-selfers. Strawson expounds on the transience view of the self. He argues that there are no persistent subjects. He can be called a panpsychist, or in his terms, a real materialist (real physicalism) and not just a physicSalist. Strawson shows that subjects have experience of themselves; they have self-experience. He claims it is the necessary and sufficient condition for having a self. There can't be any subject without subjectivity; "subjectivity" can be put in place of "subject". On his account, that means subject is an episode of subjectivity and "the existence of s (this particular episode of subjectivity) is really nothing over and above the existence of c (this particular episode of occurrent living content)" (Strawson, 2009: 414). Subject, as an episode of subjectivity, is identical to an episode of experience. Strawson endorses the identity view between experience and subjects of experience.

The real subjects for Strawson are the "thin" ones (Strawson, 2009). "Thin subjects" (Strawson argues that Descartes, Fichte, Hume, Husserl, James, and Nozick, among others, hold the "thin subject" view) are synchronic unifiers of co-conscious qualities, though not diachronic unifiers. "There's a fundamental and immovable sense in which one can't experience the self as multiple in the synchronic case" (Strawson, 2009: 90). Strawson thinks that we have short streams of consciousness. They are brief pulses of experience that can last about 2 or 3 seconds, although this is disputable. James called this temporary selves "'perishing' pulses of thought". When there is an experiential gap between them, no subject exists. Strawson thinks of subjects as real mental things. He dubs them SESMETs (short for "subject-of-experience-as-single-mental-thing"). The subject is a single, but only synchronically, for Strawson: "The unity or singleness of the (thin) subject of the total experiential field in the living moment of experience and the unity or singleness of the total experiential field are aspects of the same thing" (Strawson, 2010: 81). According to Strawson, James held a similar position on subject persistence: "Successive thinkers, numerically distinct, but all

aware of the past in the same way, form an adequate vehicle for all the experience of personal unity and sameness which we actually have" (James, 1892: 181). James reserves the word "me" for the empirical aggregate (empirical ego, the self as known), the "identity of the whole", as an objective self, and the "I" for the present, momentarily parcel of the stream, "Thought" (pure ego, the self as knower). "This me is an empirical aggregate of things objectively known. The I which knows them cannot itself be an aggregate; neither for psychological purposes need it be considered to be an unchanging metaphysical entity like the Soul, or a principle like the pure Ego, viewed as 'out of time.' It is a Thought, at each moment different from that of the last moment, but appropriative of the latter, together with all that the latter called its own" (James, 1890: 400–401). Long-term continuity is here only in the "bundle theory" sense; there are in fact many consecutive, numerically distinct selves or "Thoughts". Thin subjects are best described as: "essentially-subject-involving-experiences, briefly flaring neural synergies" (Strawson, 2009: 359).[41]

Strawson endorses the identity view between experience and subjects of experience. Every time we have a new experience with it comes a new subject. We also find the identity view in William James's *Essays in Radical Empiricism*. In the first essay, he argues for the unity of consciousness and content. James finds that experience has two different aspects, in one case it is the consciousness, as it plays the part of the knower (subject) and in the other case, it is the thing known (experience). His famous doctrine of *pure experience* can be interpreted as stating the subject/object identity, but it also relates to the subject/experience identity.

Strawson is taking into account only episodes, as if for every particular experience there is a subject of that experience. One great problem of combinationist (constitutive) panpsychist views is that if we (as macro-subjects) are made up of many

[41] Strawson's subjects eerily remind me of the descriptions of experiences of Clive Wearing, a British musicologist, who due to herpes encephalitis suffered severe amnesia, both retrograde and anterograde, stuck forever in the present that lasts between seven and thirty seconds, without a narrative self (discussed in Seth, 2021: 163–165).

subjects as parts, then we cannot say for sure which one of those subjects we really are. This is the dreaded *Problem of Self-Identification* (Roelofs, 2015: 265–304) that seems to undermine combinationism seriously. On the combinationist view, we "refer to a multitude of subjects" when we ask, "Who is talking now?" Roelofs contends. Combinationism renders self-identification impossible relative to the set of our experientially equivalent parts and probably also relative to the set of our cognitively sophisticated parts. Rather than showing how self-identification is still possible, combinationists have to bite the bullet and claim that self-identification is unimportant: knowing which set of harmoniously connected overlapping parts we belong to is all we need. (Roelofs, 2015: 303).

Strawson's theory seems to suffer from a problem of self-reference, though diachronically. Which of these subjects is me? Whether I am a human being, a human head or medulla oblongata, one might wonder in constitutive panpsychism. Similar questions could be asked of Strawson's *pearle view*.

When we talk of subject/experience identity, what experience exactly should we take into consideration? First of all, experience is holistic — the experiential field is a whole prior to its parts. Phenomenal holism holds this. Distinct experiences are "carved out" later. Holism could be defined in this manner: "Phenomenal holism — this is the view that, within a person's total psychical whole, the nature of a single identifiable experience […] is essentially determined by the other experiences occurring alongside it — synchronically — within the whole" (Basile, 2010: 107). We could rightfully ask how these thin subjects are woven into a stream of consciousness. Strawson explains: "The 'stitching software' that underwrites our sense of being a single persisting subject — and delivers a sense of the flowing continuity of experience (for those who have such experience) — is as remarked extremely powerful" (Strawson, 2009). What is it exactly that stays the same in all experiences? If there are as many thin subjects as episodes of experience, then Strawson needs to postulate some kind of phenomenal bonding relation to serve as the "stitching software", holding these subjects together diachronically. This seems like an unparsimonious posit.

Strawson's view is problematic in light of phenomenal holism. Subjects cannot be identical to single identifiable experiences. Since the synchronic experiential field as a whole is prior to its parts, there is only one holistic experience to which a subject is identical to. Dainton also points out Strawson's claim that we are identical to episodes of our experience. How do we survive sleep and unconsciousness? This is the problem of continuity (of a stream of consciousness). If Strawson is right, then we are identical to episodes of experience; we do not have experiences (Dainton, 2012: 185), and this is no ownership. Dainton would claim that overlapping chains of diachronic co-consciousness make up the stream of consciousness. The problem with the subject's persistence is how to account for the diachronic unity. This involves solving the problem of continuity of a stream of consciousness that has gaps in the form of unconscious states and dreamless sleep. For Dainton, the unity of consciousness comes from primitive inter-experiential relationships. He also denies there is something like mine-ness or non-reflective self-consciousness (Dainton, 2008: 242–243).

There is no flow of the stream of consciousness in Strawson's account; there are insurmountable gaps between short-term subject-experience-episodes. Certain worries are then raised against such an account. Since we cannot have experience of unconscious states, how can we know anything about them, even if there are such states? Zahavi asks why our phenomenal field should be fragmentary because there always seems to be some kind of "phenomenal background" of experience. Or at least there is the constant sense of mineness? Strawson's sesmet account also entails that the difference between successive "thin subjects" is as deep as between completely different selves or streams (Zahavi, 2005: 234–5). Why would all of my thin subjects be mine and not somebody else's? What makes "me" is that there is something invariant to all these sesmets in my stream of consciousness. What is the same is self-experience (in Strawson's terms). The stream of consciousness is really a stream of subjectivity, if it is streamlike at all (as James argued). The answer to these worries lies in the mineness or first-person givenness of experience.

Not just compositional panpsychists are opting for deflation. Mørch uses Strawson's identity account and expounds a hylomorphic account of causation in her emergent panpsychism. She defends a diachronic fusion account of combination.[42] She treats experientiality as "a general determinable" and reduces subjects to forms of experiential matter with the help of Strawson's identity view and Parfit's fusion. She endorses the identity view and Parfitian fusion/fission in order to make the experiential combination intelligible (Mørch, 2014: 219–220). In her view, subjects are transitory forms of fundamental experiential matter. And in the vein of Strawson's theory of sesmets it is concluded that "the subject as something that is supposed to persist through time is reduced to a series of momentary total experiential fields connected by similarity and causation" (Mørch, 2014: 216). But Strawson himself has written about equating energy with experientiality: "Energy is experientiality; that is its intrinsic nature" (Strawson, 2006: 243).

Strawson also briefly deals with the combination of sesmets (subject combination or subject-summing) when he says: "Sesmets are either single ultimates, then, or made up of a plurality of ultimates in a certain synergetic relation — if they exist" (Strawson, 2009: 295) though he gives no detailed arguments for such combination, nor does he explain the nature of this "synergetic relation". On a different occasion, he notes that he finds no problem in a plurality of subjects forming or generating a new subject. Again, there is an attempt to make subject-summing intelligible by deflating subjects themselves.

All experience is somehow bounded and unified together in the subject's phenomenal space. James explains it in the following paragraph:

> No thought even comes into direct sight of a thought in another personal consciousness than its own. Absolute insulation, irreducible pluralism, is the law. It seems as if the elementary psychic fact were not thought or this thought or that thought, but my thought, every thought being owned. Neither contemporaneity, nor proximity in space, nor similarity of quality and content

[42] See also Seager, 2010.

> are able to fuse thoughts together which are sundered by this barrier of belonging to different personal minds. The breaches between such thoughts are the most absolute breaches in nature" (James, 1890: 221).

Thoughts are always part of some mind; there is no experience of a "nobody's thought". Why is this so? How do experiences hold themselves together? Such questions increase our phenomenological need for subjectivity and experiencing subjects.

It seems that there are good reasons to take the mineness of experience as a constant and not an experientiality. In introspective observation of myself, I find a persistent sense of this "for-me-ness" of experience. Subjectivity has a self-intimating nature (Levine, 2001: 109). I think we should be appropriative of phenomenologist's insights on subjectivity and, with such knowledge, could shed some light on present matters. We should try to understand and solve problems of combination that are part of panpsychism while being self-conscious about the mineness or first-person givenness of experience. With this notion of subjectivity, as it will be argued, we can also answer the shortcomings of both Parfit's and Strawson's accounts.

4.5 UNITY OF CONSCIOUSNESS

Let us track back to the problems of boundedness and unity of subjects and make sense of them in the new light of subjectivity. I will discuss how the unity of consciousness problem would look if a phenomenological theory of subjectivity is assumed. This will show us in what way deflation could be wrong. Chalmers and Bayne (Chalmers & Bayne, 2003) define *The Unity Thesis* as follows: "Necessarily, any set of conscious states of a subject at a time is unified". To answer the boundary problem is to answer the unity problem; there is a deep connection between these issues. They are not the same but seem to be in a close relation, because notions of unity and boundedness are close. In terms of phenomenology, I regard the subjectivity of experiences as what sets the boundaries of subjects. Both the unity and boundedness

are explained by the metaphysical fact that they belong to the same bearer, the same subject that has them, and mineness, first-person givenness and pre-reflective self-awareness point to this. One proposed solution to the subject-summing problem is the phenomenal bonding relation strategy (Goff, 2009), positing a special kind of relation between subjects that seems to unite subjects into a composite subject, though the bond is unknown to us. It is such because we can only introspect within a subject. Perhaps the intrinsic nature of physical relations is the phenomenal bonding relation.

Although phenomenal bonding is an intersubjective relation, it is often framed as a problem of intrasubjective relations. Chalmers (Chalmers, 2016) claims that phenomenal bonding is "co-consciousness", i.e. the relation of the unity of consciousness. But what the co-consciousness relation really is? In itself, it is empty, undefined. This notion doesn't seem to explain much, it just states that some phenomenal states are experienced together, conscious together, and phenomenally unified. And why are they experienced together? James writes on the co-consciousness relation:

> The conjunctive relation that has given most trouble to philosophy is the co-conscious transition, so to call it, by which one experience passes into another when both belong to the same self. About the facts, there is no question. My experiences and your experiences are 'with' each other in various external ways, but mine pass into mine, and yours pass into yours in a way in which yours and mine never pass into one another. (James, 1912: 47)

Different selves are related in various ways through external space, but experiences are "with each other" in the inner space of the self. Chalmers has pointed out that there is a question of whether the co-consciousness relation is transitive or not. Dainton (Dainton, 2011) imagined how a non-transitive view of co-consciousness could make the combination problem coherent. But it has to be the case that co-consciousness is transitive, and all experiences (states) are co-conscious in a total state of consciousness of a subject. That they belong to one subject tells us where the transitivity stops, so to speak. It shows where the

boundary of consciousness is. Just look at the James' paragraph; it states that experiences are co-conscious "when both belong to the same self", not the other way round. Bayne and Chalmers argue that the unity thesis cannot be explained by starting from "our concept of a subject". Their argument does not succeed because it assumes the bundle theory of the self, which is not the only available option on the metaphysical "table".

So, how to explain the phenomenal unity? One possibility that is worth exploring is that self-consciousness accounts for the unity of consciousness (Bayne, 2004). Bayne explores renditions of unity based on self-consciousness. The psychological constraint on co-consciousness states that "experiences can be co-conscious only if the subject of those experiences is aware of them as their own" (Bayne, 2004: 229). The robust account of the unity of consciousness of this sort would take that self-consciousness explains the unity of consciousness, though Bayne doesn't defend such a strong account; he just claims that self-consciousness constrains the unity. I think that a more robust account can and should be argued for. Bayne tries to argue against the psychological constraint based on the considerations of thought-insertion, depersonalisation and Cotard delusion cases. These patients have a phenomenally unified perspective but lack ownership ("the bare sense of being the subject of an experience" in Bayne's terminology). So any defender of the self-consciousness account of unity will have to show that in these cases, a sense of ownership is preserved.

One could say that experiences are unified in the self as in a kind of space ("subjective space"). Talk of the field-like characteristics of subjectivity is not new, e.g. a "field of first-personal givenness of experience" (Zahavi). Subjectivity is the foundation of experience, the space where experience is manifested, and the relation of the subject to the experience could be the same as spacetime is related to its material objects. Therefore, the "spatial relations" between experiences are just relations of the subjectivity space. Experiences are unified because they belong to the same space of subjectivity. Their co-consciousness is grounded in their shared subjectivity, i.e. the same first-personal givenness. This is why the co-consciousness relation may be misleading.

4.6 BACK TO DEFLATION

Since most phenomenologists endorsed the existence of pre-reflective self-consciousness, we should appraise such theories of the self. Phenomenology is too broad; there are too many phenomenological theories of the self to consider them all in the course of this chapter. I will concern myself with more recent phenomenology-inspired theories and those that criticise phenomenological theories, though they are similar in spirit. To make the contrast clear, I will consider egological and non-egological theories. In any case, it is of paramount significance that panpsychists consider theories that are serious about subjectivity (namely, pre-reflective self-consciousness). To argue against deflation of the subject/ego, we need to understand what the subject is and so examine other theories of the self. What all these authors agree on in their theories is that there is an abiding dimension of giveness, presence, witnessing or openness: "field of the first-personal givenness of experience" (Zahavi), "dimension of first-personal manifestation of the experiences" (Fasching).

Those who defend panpsychism should consider phenomenological theories of subjectivity and acknowledge that there is ubiquitous pre-reflective self-consciousness and that there is an experiencing subject that it points to. If there is such awareness, then the subject combination problem and the unity problem are to be resolved according to that fact. Pre-reflective self-consciousness explains why the subject combination problem is intractable and how we could solve it, but also why the unity of experience is based on the experiencing subject as their substratum or bearer. Perhaps different inferences on these panpsychist problems could be reached depending on the different understanding of pre-reflective self-consciousness. It would depend on whether one maintains an egological or non-egological theory of self-consciousness. Some philosophers of subjectivity who are influenced by Phenomenological tradition see this self-awareness as individuating and pointing to the existence of an ego (egological theories). Endorsing such a view of self-awareness is more likely to lead to the conclusion that there is no plausibility in subject combination and that unity is to be explained by the presence

of the ego. Others (The Heidelberg School, Sartre and Gurwitch) argue for non-egological theories of self-consciousness and for the anonymity thesis. Taking up such a stance on pre-reflective self-consciousness could prove to be more compatible with the deflationist position in panpsychism. Panpsychists should consider this and base their solutions to the aforementioned problems on these phenomenological theories. Panpsychists should not ignore the importance of self-consciousness if they are to construct good theories of consciousness.

Resolution of panpsychist metaphysical problems depends on how they understand the concept of the experiencing subject (deflationary/non-deflationary), and phenomenology has a lot to say about the "subject". There is something like pre-reflective self-consciousness, but it can be understood in different ways (egological/non-egological), so this, too, has repercussions for panpsychist theories.

Suppose there is one individual mineness or pre-reflective self-awareness on a subject or pointing out that there is one subject in question. In that case, the combination of such subjects seems less plausible. If the dimension of mineness is anonymous (Fasching), self-awareness does not point to any individual subject (non-egological theory), and there are only experiences connected in a bundle by co-consciousness relation, the plausibility of subject combination (subject-summing) is increased. Therefore, all this has important repercussions in the subject combination debate among panpsychists. Philosophers of self-awareness (Guillot, Siewert, Nida-Rümelin) have argued that such properties as mineness or me-ness (and pre-reflective self-awareness) point to a relation between an experience and a subject of experience, and this subject seems to be the same one in many synchronic and diachronic experiences. If it is to be judged by the phenomenology of self-awareness and mineness, the deflation of subjects proves to be an invalid strategy that cannot be justified. The accounts of Zahavi, Strawson and Fasching, though not panpsychist, are also deflationary, but as it was argued, the phenomenology of pre-reflective self-awareness seems to point to a more inflationary position when it comes to the nature of the self. Zahavi and Strawson, in the end, present precarious positions.

The phenomenology of self-awareness can give us introspective knowledge about the nature of subjects, and this, in turn, would have important consequences for the plausibility of constitutive panpsychism and for deciding on the possible solutions to the subject combination (subject-summing) problem. The route to subject-summing is indirect but valuable. Namely, if pre-reflective self-awareness shows us that we are individual substances, deflation of the subject would not look very promising as a strategy of a would-be panpsychist for solving combination problems. Moreover, if deflation is not plausible, then the subject-summing's viability is also questioned. In that case, the main strategy that paves the way for subject combination in panpsychism is also brought down.

The aim of this chapter was to show that panpsychists should include more considerations on pre-reflective self-awareness and mineness (subjective character of consciousness in general) in their future metaphysical theories. Even if one is not persuaded that the self is a substance of some sort, there is an aspect or a dimension of mineness that needs to be reckoned with by panpsychist theories of consciousness, especially when it comes to the problem of subject combination. Ultimately, this could challenge the commonly assumed reaction strategy to the combination problem — deflation of the subject. The panpsychists seeking a way to overcome the subject combination problem would benefit from exploring the phenomenological theories of selfhood and pre-reflective self-awareness.[43]

4.7 SUBJECTS EMERGE!

The combination of subjects poses a seemingly insurmountable challenge for constitutive panpsychism. Moreover, it appears inherently unsolvable. Goff contends that the mere presence of a set of subjects does not necessarily imply the existence of another subject a priori. The mechanism of this occurrence remains elusive, although there is a potential, albeit unidentified

[43] I discussed these problems in Nešić, 2017.

avenue for it to transpire. Coleman, while acknowledging Goff's argument, maintains that combining subjects to generate new subjects is metaphysically impossible. This stance eliminates constitutive panpsychism, as it implies a form of brute emergence — an outcome undesirable to panpsychists who sought to avoid such phenomena from the outset.

Coleman rejects the plausibility of constitutive panpsychism and advocates for panpsychists to adopt a neutral monist perspective. While I find Coleman's robust "no-summing" argument compelling, I diverge from his conclusions and propose solution to the predicament. I believe that Coleman's argument effectively challenges constitutive panpsychism, but does not necessarily negate alternative forms of emergent panpsychism.

Strong, radical emergence becomes pertinent if one endorses fusion (Humphreys, 1997) to derive macro-subjects from micro-subjects. Seager and Mørch both subscribe to variations of this emergence in their theories. In the case of fusion, lower-level subjects give rise to another subject while ceasing to exist themselves — essentially being annihilated in the process. Mørch has suggested that subjects may partially survive, undergoing change within the whole to which they belong but persisting in some form, especially evident in the absence of fusion in the brain. However, this account lacks a clear-cut emergence of a novel macro-subject in the strong sense, as it merely comprises the micro-subjects that constitute it, albeit altered to some extent. This absence of genuine combination in fusion aligns with the panpsychist's concerns, as only radical, brute emergence would occur, rendering it unhelpful for constitutive panpsychism. Mørch's position, while avoiding strong emergence, is deflationary, since it lacks a macro-subject in any robust sense.

Emergent panpsychism that I am considering here is one in which the micro-subject *is* the macro-subject (or, at least, some aspect survives and is identical), so there is no combination whatsoever (and no annihilation). To avoid the pitfalls of constitutive Russellian monism, while staying true to the indivisible nature of subjects, I propose that panpsychists should consider embracing emergentist Russellian Monism. The Russellian monism view I would like to discuss now, uses Shoemakerian emergence to

obviate the need to deny the micro-physical closure and is a more plausible and desirable form of emergentist panpsychism.

4.7.1 *Shoemaker on Emergence*

Shoemaker has formulated a captivating and innovative theory, drawing inspiration from Broad's concept of emergence as outlined in "Mind and Its Place in Nature" (Broad, 1925). Broad illustrates this idea using the example of silver-chloride and its properties in relation to its chemical elements, silver and chlorine. Shoemaker's propositions seem to echo Broad's insights:

Broad contends that certain properties, termed "emergent properties," cannot be deduced from combinations of other properties. He suggests that there are latent properties in substances, only becoming apparent when these substances are combined in specific ways. This dual description includes novel emergent properties and those manifest when there is no combination, referred to by Shoemaker as "emergent engendering" ways (Broad, 1925: 66).[44]

Shoemaker introduces the concept implying that when micro-entities are combined in an emergent engendering way, they possess two types of micro-structural properties. The first type, called Type-1 properties, are entirely specified by the micro-manifest powers of the constituent micro-entities and their relationships. The second, Type-2 emergent properties, are characterized as being specified in terms of all the powers, both micro-latent and micro-manifest, of the constituent micro-entities (Shoemaker, 2002: 56).

According to Shoemaker, Type-2 properties supervene on the Type-1 properties. In an emergentist framework, macro-entities' properties are not predictable based solely on micro-facts; rather, they are realized in emergent micro-structural properties. The micro-facts determine the macro-facts, but in this emergent perspective, the micro-facts include the instantiation of micro-latent powers. In this view, there is no need for macro-properties

[44] See Broad, 1925.

to be realized separately; the emergent properties embody those macro-properties. Shoemaker's theory explores the intricate relationship between Type-1 and Type-2 properties, shedding light on the dynamics of emergent phenomena.

4.7.2 *Shrader's Reworking*

On Shoemaker's account, the dependence relation between emergent and physical properties is nomological supervenience (Shrader, 2010: 286). Shrader has criticized that Shoemaker's account fails to meet the conditions of *Minimal Ontological Emergence,* which states that emergent properties are dependent on, but not reducible to the physical properties, and that they make a novel causal difference. He also tried to resolve some of the problems of Shoemaker's account in the same paper.

There is a problem with Shoemaker's emergence, namely that emergent properties are structural, something that is commonly argued against. For example, O'Connor (O'Connor, 1994: 94) says that structural properties cannot be involved in downward causation and are not causally efficacious. But there is a problem with this reasoning, and Shrader shows it. This supposes there are only micro-manifest powers, and Shoemaker's account avoids this by adding micro-latent powers. Shrader gives an example of such powers. Micro-manifest power could be the power of bestowing a mass; it is always manifest. The complete micro-physical theory would mention these powers but not the latent micro-powers, such as the "power of bestowing the property of being in pain" (Shrader, 2010: 291). This is because they are manifest only in higher-level entities.

Shrader then marks a huge problem for Shoemaker's emergence: every type-2 property seems to be nomologically equivalent to a type-1 property (Shrader, 2010: 294). And if a causal theory[45] of properties is endorsed, and Shoemaker does endorse it,

[45] It seems that even Shoemaker himself doesn't hold that all there is to a property is its causal role. As Pereboom states, Shoemaker holds the view that properties "also feature intrinsic aptnesses for the causal roles that individuate them" (Pereboom, 2015: 312).

then nomological equivalence implies *identity*. So, these are not emergent properties at all. Shrader states that even if the causal theory of properties is abandoned and you hold that the powers contributed by properties vary across worlds, this still fails to be a genuine ontological emergence. Shrader changed Shoemaker's account to avoid these. Type-2 property can be viewed as a conjunctive property, whose conjuncts are type-1 property and a type-3 property, "the property of having such and such constituents possessing such and such micro-latent causal powers and related in an emergence engendering way" (Shrader, 2010: 295). Type-3 are the real emergent properties, the properties of having micro-latent powers. This eliminates the problem of type-3 and type-1 properties being nomologically coextensive and identical.

Furthermore, Shrader points out that both Shoemaker's and the new emergence imply the denial of physical causal closure. Shrader solves this by arguing that the synchronic dependence of type-3 on type-1 properties should be regarded as *causal* rather than mere nomological supervenience (Shrader, 2010: 297). It should be a case of *causal dependence*. Then, there would be no breaching of the causal closure since type-1, type-3, and some type-1* properties would be in the same causal chain. There is a denial of causal exclusion, but with "an explanation". Shrader does not say it is micro-physical causal closure, though it seems that is the case, since these are all micro-structural properties. All this seems to build a coherent picture of emergence.[46]

[46] Hedda Hassel Mørch and Luke Roelofs have commented that the problem with this could be that there is no evidence of micro-latent powers manifesting at the higher levels; they are not discovered by physics and are excluded by micro-physical closure. Just because physics can't discover micro-latent powers, that doesn't mean they don't exist or cannot exist already at lower levels; they can be detected, but only at higher levels, when they become manifest. These are *metaphysical* problems, and Shoemaker's account can help us solve or avoid them. I am not sure if properties are latent in such a way and if there is empirical evidence for this, but when it comes to the problems of panpsychist subject-summing, this account seems to me to be a very useful way of thinking about subjects. There are such micro-latent powers as Shoemaker thinks when it comes to the micro-causal closure of the problem. Physics knows only of micro-manifest powers, but these latent powers still have causal efficacy, only that they become manifest when the parts are combined in special ways. So their effects are detectable higher, but does this mean they can't be fundamental (working on the micro-physical level)? This would be a form of strong emergentism, but the macroscopic whole is the *loci* of emergence. Perhaps we could

4.7.3 *Emergence in Panpsychism*

Let us apply this revised concept of emergence to panpsychism. In the assumed Russellian identity theory, micro-powers are posited as identical to micro-subjects. This holds significance because the intrinsic character, serving as the explanatory factor and bearing causal relevance, is the focal point of this theory. If we were to adopt Shoemaker's account without alterations, we might end up with a form of identity panpsychism. By drawing an analogy, we could identify micro-subjects with type-1 properties and macro-subjects with type-2 properties, which are emergent. However, type-2 properties are combinations of both micro-latent and micro-manifest power, and it seems reasonable to argue that the micro-latent powers truly emerge, aligning with the desired characteristics for macro-subjects. Conceptualizing macro-subjects (o-subjects) as Shoemaker's type-2 properties appears to introduce a new combination problem.

A more precise approach would be to seek macro-subjects that exhibit properties specified entirely in terms of micro-latent powers, as per Shoemaker's terminology. To avoid the identity of type-2 and type-1 properties, thus preserving the concept of emergence, Shrader introduces the L property as type-3, which is considered emergent.

According to Shrader, type-3 properties are not nomologically equivalent to type-1 properties. It can be argued that type-3 properties nomologically supervene on type-1 properties, implying a dependence relationship. However, Shrader finds this dependence problematic and deems it preferable to interpret it as causal dependence. In this view, there is intelligible causation between them. In the context of Russellian Monism (RM), this implies a causal dependence between o-subjects and micro-subjects.

When we apply reworked Shoemaker's account on RM, o-experience is emergent, so it should be equated with the

use the answer Mørch gives in her dissertation (see pages 206–210), citing Cartwright (Cartwright, 1994: 281), that even if there is a strong emergence with latent micro-powers, this does not entail that the laws of physics are violated. Cartwright argues that reality is a patchwork of laws; metaphysical nomological pluralism.

L properties (type-3) (having micro-latent powers) and micro experience is to be equated with the having of micro-manifest powers. Endorsing the causal dependence relation between these properties obviates the need to deny micro-causal closure. There seems to be no problem in Shoemaker/Shrader's emergence with micro-physical closure because o-experience is not something over and above micro-experience, *it is* from the fundamental micro-level.

4.7.4 *Subjectivity Emergence*

Goff's term *o-subject* has left a possibility open, the possibility that the pre-theoretical subject is identical to a micro-subject, but it seems implausible to identify *me or you*, as a subject of experience, with some "micro-level entity". If micro-subjects are the same as particles, how could one particle cause the behaviour of my pre-theoretical mind? This could lead to epiphenomenalism.

This position seems to be the one Chalmers takes into consideration as a possible reaction to the combination problem. One option is to hold that macro-subjects are identical to certain micro-subjects (Chalmers, 2015: 270).[47] Solution is akin to Leibniz's "dominant monad" view. The problem is how a human subject is identical to a single fundamental particle in the brain and how does it has complex phenomenal properties (all this on Russellian monist view). To give one possible answer: it doesn't need complex properties; it just needs one for subjectivity. Subjectivity is simple. It could just need a property of being a point of view.[48]

Since we tend to think of subjects as non-spatial, this doesn't seem so far off. Does "size" really matter when it comes to subjects? I don't understand what would saying that o-subjects are macro- instead of micro- really amount to, in case there is no constitution and subjects can't combine. What would really be the difference? The macro-physical object is composed of many

[47] The others are giving up constitutive panpsychism for emergent panpsychism, and quantum entanglement deflating the subject.

[48] This is where deflation of subjects would be helpful.

micro-physical objects, but macro-subjects don't have to be composed of many micro-subjects. If we take a subject of experience to be indivisible, there is no difference between a micro-subject and a macro-subject. If they are to combine into macro-subjects, does that mean that micro-subjects are spatially related in some way? However, the problem of mental-physical isomorphism still stands for the Russellian monism version of panpsychism.

One problem with Shoemaker's emergence is that latent dispositions are situated on the micro-level, and this doesn't make downward causation coherent. That is why O'Connor (O'Connor, 1994) thinks emergent dispositions should be situated in the macro-whole itself and so they have a downward influence on the behaviour of involving parts. Shoemaker envisions in his account that when micro-latent dispositions become manifest, they are not located on the macro level, but they *do have large-scale effects*, so no epiphenomenalism should entail. Another problem that needs to be answered involves making sense of the dormancy of subjects of experience. Powers are latent and manifest, not the objects that have those powers. How might we solve this? Such a predicament, perhaps, could be avoided if it is considered that this is the Russellian identity theory framework, and I already suggested it is the most plausible form of Russellian monism. If a phenomenal property can be identical to a dispositional property, then it seems less implausible to think of *any property of subjectivity* as having the characteristics of a power. It should be noted that everything pertains to the properties of subjects, whether experiential or dispositional, since it is unclear how an entity could be "latent".

How can we comprehend the differentiation between *latent* and *manifest* aspects in subjects? Dainton's C-system theory is the sole theory that connects potentiality to subjects of experience, specifically through capacities for experience. This is articulated as *The Potentially Conscious Self* thesis (Dainton, 2008). One way to grasp the emergence of subjects is through their transition from potentiality to manifestation or actuality. Dainton rejects the notion that subjects are inherently conscious entities (*Essentially Conscious Self* thesis), contending instead that they are entities with the potential to be conscious. The subject

or self is considered potentially conscious because it possesses the *capacity for consciousness*. However, a potential issue arises as this definition of the subject becomes dispositional, deviating from the standard conceptualization of subjects. Therefore, both Shoemaker's subject emergence account and Dainton's potentially conscious self thesis require an explanation of how subjects of experience can be dispositionally defined. These capacities may lie dormant and, at certain times, contribute to a unified stream of consciousness. Combining such a view of the subject with the present account could prove beneficial. While there are profound differences between these accounts, one being Lockean and the other Leibnizian; this example illustrates a plausible approach to making the claim intelligible.

Concerning the property of subjectivity, if it is indeed a property, could subjectivity be latent and manifest? The primary rationale against the divisibility of subjects (as seen in fission/fusion of subjects) is the nature of subjectivity itself. The concept of subjectivity is inconceivable regarding division, cutting, or fusion. The term "experiential combination" is employed because it is conceivable to fuse and fission experiences and streams of consciousness. One could imagine scenarios, such as split-brain (split-stream) cases, where experiences are rearranged or transferred. However, the same does not hold for subjectivity; sharing someone's point of view is unimaginable. It is conceivable to divide a river, a stream of experiences, but can subjectivity be divided? Carrying Coleman's argument to its extreme implies the absence of any form of subject combination, not even in cases of emergence (fusion cases). According to this perspective, a subject cannot be emergently caused or produced from other subjects.

There are two possibilities for Coleman: that subjectivity is a structural property of the micro-subjects combining or there is a strong emergence as the causal production of a macro-subject from micro-subjects. On the other hand, there are other possible ways of understanding emergence. I think the best option is to go with Shoemaker's emergence. Coleman entertained the thought that one point of view could survive to become a macro-subject.[49]

[49] See Coleman, 2013.

A macro-subject is just one of the micro-subjects (dominant monad). This is not a combination anymore, but emergentism, only of a different kind, the way Shoemaker understood emergence.

The same goes for the deflationary views on subjects that are sometimes invoked to defend experiential combinations. Even in deflated (thin) subject accounts, this seems implausible. Thin subjects are still strong units, be they synchronic or diachronic. This is especially visible in thin subject positions that stress subjectivity or mineness (first-person givenness), like Zahavi and Strawson. Some want to argue for reducing subjects to experiences because, in such a case, combination and causal production would seem more likely to occur. The argument seems so strong that neither constitution nor emergence (causation) will work. In the end, there is no combination whatsoever. Because of the indivisibility of subjects, the only alternatives are radical emergence and that one of the micro-subjects becomes the macro-subject, one possibility being that a micro-subject has been dormant and becomes active on higher levels (the "dormant monad" view). The other one is a novel solution from Shoemakerian emergence, but both are dominant monad views.

We get an account of non-radical emergence; there is no becoming of new subjects out of nothing, and it vouches there is no breaching of the microphysical closure because all the causal influence drains at the bottom. This harkens us back to the problem with causation, so panpsychist should bite the bullet and claim, as the Shoemaker account has entailed already, that the relation is *identity*. This would lead us straight into a Leibnizian solution. If we don't want to postulate unknown latent micro-powers/subjects, the same conclusion could be secured if the subset realization account is endorsed, so that the macro-subject is just one of the micro-subjects as a part of the micro-realization whole.[50] And this is still very much in the spirit of Leibniz's ideas that the material body is composed of many monads. The realization account can potentially solve the epiphenomenalist threat;

[50] The causal efficacy comes from the whole, but it is attributed to the part. It would also be useful to combine this account with the Russellian realization theory, to make it intelligible

it is the whole of many monads (micro-subjects) that exerts the causal effect, though it is attributed to a part of it, to one monad.

Some remarks from Basile are informative. Basile argues that experiential combination leads to a contradiction. There are two requirements that a viable panpsychist theory should satisfy: "[A] The panpsychist should reject the idea that larger experiential wholes are brought into existence by way of simple addition of lesser experiential realities; [B] The panpsychist should fully acknowledge that experiences are private. This means that the notion of mental composition should not be construed as involving the idea that experiences can be literally 'shared' — 'owned' as it were by two different subjects." (Basile, 2010: 111). What known panpsychist theory could answer them? Basile answers that Leibniz's theory of causally independent monads is the one, and that the mind is an *indivisible substance*. In the end, Basile asks if a recurrence of a Leibnizian metaphysics would be a price too high to pay for a panpsychist? Though he does not answer this question, given what was discussed in this chapter and the account defended here, I would say that it is a route that shows promise for a panpsychist.

I think those who take Coleman's anti-summing argument seriously and want to stay on the ground of panpsychism are forced to take the plunge and endorse the conclusion that subjects are fundamental entities by taking up one of the positions discussed so far. Then, instead of the combination problem, the issue of epiphenomenality should be addressed. Benefits would be a strong notion of emergent o-subjects and no breaching of the micro-physical closure, though with little elegance. So, for this to be a viable alternative for panpsychists as a way to avoid the combination problem and not slip into neutral monism, the problem of epiphenomenalism should be met. Another large problem for this solution that needs working out is making sense of the alleged latency of subjects of experience.[51]

[51] I developed this in Nešić 2018b. Hibbert has criticized my account and offered a similar one in which "the dominant instantiation of phenomenality 'spreads itself out' to overlap the experiential content in the system" (Hibbert, 2022: 253). He argues that resolving the combination problem in constitutive micropsychism involves creating a mechanism wherein a particular manifestation of

We could, perhaps, strengthen the case for emergent subjects if we back up the Shoemaker's emergence with O'Connor and Jacob's account of emergence. But this would come with the price of endorsing a bundle view of subjects. We could say that the emergent powers together make the *emergent subject* or the *emergent individual* (O'Connor & Jacobs, 2003). Emergent powers of the Shoemaker's account are on the same micro-level, but they pertain to the whole and are holistically causally efficient, though there is no downward causation. If we keep the Shoemaker account in the game, o-subject that the emergent (latent) powers make together will be on the micro-level, though its effects are macro and holistic and not epiphenomenal. This o-subject is made of all the latent powers that are manifest and cannot be epiphenomenal. To avoid the epiphenomenality of micro-subjects, we could say only one subject emerges, the one that all the emergent powers make together (in a bundle). O'Connor and Jacobs take the *trope bundle* view of Peter Simons as an example of how to get a new, emergent object/subject/individual, though they do not adopt it. In our case, a bundle or a nucleus of emergent tropes constitutes an emergent individual, a new substance or a new subject of experience. They think that the emergence of properties leads to a stronger claim that there is an emergence of individuals. Preserving the Shoemaker account also keeps the micro-physical closure, denied in O'Connor and Jacobs' strong version of emergence. And in a panpsychist framework, in the mental realm, these emergent individuals are *emergent subjects*.

phenomenality takes precedence over other manifestations. This dominance of phenomenality, denoted as 'Z,' should be structured in a manner that allows it to coincide with and encompass other manifestations of phenomenality, namely 'Y' and 'X.' (see figures in *Non-fused Dominant Phenomenal Overlap* chapter of his dissertation and discussion of my view).

5. SELVES AS PATTERNS OF INFORMATION

> It was not exceptional at that age for me to feel moments of complete disconnection, periods of total self-absorption — studying closely the lines on the palms of my hands or watching my shifting shadow as I leaned backwards and forwards in slow and rhythmic movements. But this was something else, an experience unlike any other, as though the room around me was pulling away from me on all sides and the light inside it leaking out and the flow of time itself coagulated and stretched out into a single lingering moment. I did not and could not have known it then, but I was having a massive epileptic seizure.
>
> *Born on a Blue Day* (Tammet, 2006: 36)

5.1 INFORMATION INTEGRATION

In this chapter I will investigate the status of subjectivity in Integrated Information Theory. This will lead me to examine whether Integrated Information Theory can answer the hard problem of consciousness. In itself, Integrated Information Theory does not seem to constitute an answer to the hard problem but could be combined with panpsychism to yield a more satisfying theory of consciousness. I will show that even if Integrated Information Theory employs the metaphysical machinery of panpsychism, Integrated Information would still suffer from a different problem, unable to account for consciousness's subjective character.

Integrated Information Theory (IIT), proposed by Tononi and colleagues (Tononi, 2008; Oizumi et al., 2014; Tononi & Koch, 2015), is a neuroscientific theory designed to address fundamental questions about the nature of consciousness. It aims to determine which physical systems possess consciousness and

what these systems are conscious of. The theory also delves into quantifying the level of consciousness within a system. This exploration has practical implications, particularly in understanding borderline conscious states in patients with brain injuries. The corollaries of Integrated Information Theory suggest that aggregates and machines are not conscious entities, while unexpected entities like photodiodes might exhibit consciousness, albeit in a quantitatively and qualitatively minimal form. This theory has been hotly debated for some time and has even been dismissed as pseudoscience in a recent letter signed by a number of researchers (Fleming et al., 2023).

The first section of this chapter introduces the axioms of Integrated Information Theory, which will be further explored. In the second section, the discussion centres on whether IIT can address the hard problem of consciousness (Chalmers, 1996). One potential avenue is interpreting IIT as a form of panpsychism, as suggested by some philosophers. This interpretation could provide the necessary metaphysical framework to tackle the hard problem, and Section 3 examines whether IIT can be viewed as an emergentist panpsychist theory of consciousness. Section 4 argues that even if IIT is interpreted as panpsychist, it faces an additional problem — failing to account for the subjectivity of consciousness. Various versions of panpsychism that attempt to augment IIT also grapple with this issue. To elaborate on this shortcoming, the chapter draws on notions of mineness and pre-reflective self-awareness from recent works by authors such as Zahavi, Kriegel, and Nida-Rumelin. The chapter's modest aim is to investigate the treatment of subjectivity, specifically the subjective character of consciousness, within Integrated Information Theory, assessing whether IIT adequately addresses subjectivity, or considers it at all in its theory of consciousness.

Let us put forth some essentials of Integrated Information Theory. Adherents of IIT start by taking consciousness seriously and defending axioms and postulates concerning it. Axioms are the essential phenomenological properties of consciousness. These are taken by authors of IIT to be "immediately evident". This in itself can be very problematic, but let us take the axioms as they stand in IIT and see where that would take us. The five

axioms of Information Integration are *existence, composition, information, integration and exclusion* (Oizumi *et al.,* 2014; Tononi & Koch, 2015). The first axiom, that of *intrinsic existence* (Tononi & Koch, 2015), tells us that *consciousness exists*. This is unquestionable. IIT is a theory that respects phenomenology and takes the existence of consciousness as a fact.

The second axiom, *composition*, appeals to experiential combination: "Consciousness is compositional (structured): each experience consists of multiple aspects in various combinations." (Oizumi *et al.,* 2014: 2). Our phenomenal field is made up of many experiences combined together. Does holism come first and out of the unity of consciousness, we "carve" out the many distinct experiences later? We will address this question later when the problem of the unity of consciousness is analysed.

The third axiom, *information*, claims that consciousness is *informative*: "each experience differs in its particular way from other possible experiences" (Oizumi *et al.,* 2014: 2). Then, with the fourth axiom, *integration*, comes holism again: "Consciousness is integrated: each experience is (strongly) irreducible to non-interdependent components." (Oizumi *et al.,* 2014: 3). So, "experience" is also something over and above its components. Observe that this seems to go against what was said in the second axiom since composition is nothing over and above the sum of its parts. *Integration* is the most important axiom because it grounds the central idea of IIT in the first place. Finally, there is the axiom of *exclusion* that states: "Consciousness is *exclusive*: each experience excludes all others — at any given time there is only one experience having its full content, rather than a superposition of multiple partial experiences; each experience has definite borders — certain things can be experienced and others cannot; each experience has a particular spatial and temporal grain — it flows at a particular speed, and it has a certain resolution such that some distinctions are possible and finer or coarser distinctions are not" (Oizumi *et al.,* 2014: 3). *Exclusion* contains a *boundary* principle and an implication that no conscious subject's parts are themselves conscious.

These phenomenological truths are mirrored in *postulates* of the physical substrates that realize phenomenology. Postulates

are the properties that physical mechanisms must have in order to support consciousness, and identity is the relation between the phenomenal and the physical in IIT. The names of the postulates are the same as the names of the axioms, but they pertain to cause-effect structures since causation as information is the "shadow" aspect of experience. For IIT, Tononi has taken Bateson's position on information: "Information is a difference that makes a difference" (Bateson, 1972).

The authors state in the first postulate that a system of mechanisms must exist *intrinsically*. That means that it has the cause-effect power independent of the extrinsic factors, "power upon itself". This is one pointer in the direction of *subjectivity* as an aspect of consciousness and towards an account of it that we find in IIT, though from the point of a system of mechanisms. Every system is structured; it has subsystems. The system has a particular cause–effect structure, given the information postulate. The system is intrinsically *irreducible* (unified), and the cause-effect structure is definite. The structure that is maximally irreducible intrinsically is a conceptual structure made of maximally irreducible cause–effect concepts, as per *Exclusion* (Tononi & Koch, 2015: 7).

Regarding the mechanisms that support or are the correlates of consciousness, the authors say: "Integrated information is information that is generated by the whole mechanism above and beyond the information generated by its parts. This means that, with respect to information, the mechanism is irreducible." (Oizumi *et al.,* 2014: 7). Tononi and Koch also make the following claim of identity:

> Identity: an experience is identical to a conceptual structure that is maximally irreducible intrinsically (MICS, a constellation of concepts in qualia space)... a conceptual structure completely specifies both the quantity and the quality of experience: *how much* the system exists — the quantity or level of consciousness — is measured by its Φ^{max} value — the intrinsic irreducibility of the conceptual structure; *which way* it exists — the quality or content of consciousness — is specified by the shape of the conceptual structure. (Tononi & Koch, 2015: 9).

Therefore, IIT should be able to measure the quantity and quality of experience in a system and the level of consciousness in a system. Even if this is true, should we conclude that IIT fairly describes what it means to be conscious? This has to do with answering the hard problem of consciousness and, ultimately, with phenomenology and subjectivity of consciousness.

5.2 PROBLEMS OF CONSCIOUSNESS?

Aaronson has criticized IIT, saying that it does not solve the hard problem. He argues that it is also unable to answer "the pretty-hard" problem of consciousness. Aaronson defines "the pretty-hard" problem as a problem: "How to construct a theory that tells us which physical systems are conscious and which aren't — giving answers that agree with 'common sense' whenever the latter renders a verdict — is one of the deepest, most fascinating problems in all of science" (Aaronson, 2014). To respond to this worry, a theory of consciousness needs to tell us which system is conscious, and this has to agree with our common sense intuitions about what conscious entities are. Since IIT posits consciousness in all sorts of counter-intuitive places and entities (like photodiodes), it does not even constitute an answer to the pretty hard problem. Chalmers has analysed the pretty-hard problem into several distinct problems and claims that IIT is a partial answer to at least one of them: "Construct a theory that tells us, for any given physical system, which states of consciousness are associated with that system".[52] Aaronson goes on to argue that although a large Phi value (that is central to IIT) may be a necessary condition for physical systems to be conscious, it is still not a sufficient condition for consciousness.

From the quoted Chalmers' passage, it would seem that Information Integration is a part of the answer to the *easy* questions of consciousness: "When we think and perceive, there is a whir of information-processing, but there is also a subjective aspect. As Nagel (Nagel, 1974) has put it, there is something it is like

[52] See Mindt, 2014 for discussion.

to be a conscious organism." (Chalmers, 1995: 3). It appears that IIT doesn't really approach the hard problem after all. IIT seems to lack the metaphysical "heavy machinery" to answer this problem. Something more is needed for IIT to constitute a valid solution to the hard problem.

Others have already pointed out this problem with IIT and speculated that merging IIT with panpsychism would be beneficial (the other alternative is Russellian monist panpsychism). It was pointed out that the authors of IIT had already argued that the theory entailed a form of panpsychism. Panpsychism would help IIT solve the hard problem of consciousness, and IIT would help panpsychism solve its distinctive *problem of combination*. In this chapter, I will not consider other possible interpretations of ITT, like the one that understands IIT as more akin to the *dual-aspect* theory of Chalmers (Chalmers, 1996), but just concern myself with the panpsychist interpretation.

Since Integrated Information Theory is rather vague about many important concepts pertaining to the whole issue of consciousness to the extent that it may be questionable what it is a *theory of*[53], it may prove beneficial to improve this theory with certain metaphysical clarification. More on integration of Integrated Information Theory with panpsychism will be said in the next section. Even though merging of panpsychism with IIT could prove important for both positions, I will point out a new problem for the improved Integrated Information.

5.3 PANPSYCHIST INFORMATION

Is there a reason to interpret IIT as a panpsychist theory?[54] How could one corroborate such a thesis? It was Chalmers who first argued for the deep connection of information and experience: "If there is experience associated with thermostats, there

[53] As Ned Block remarked at one point to Tononi: "You have a theory of something, I am just not sure what it is". See Cerullo 2015.

[54] Hohwy interprets IIT as *infopsychism*, a "theory-driven, naturalistic view" in which consciousness is ubiquitous, since integrated information is ubiquitous (Hohwy, 2022).

is probably experience everywhere: wherever there is a causal interaction, there is information, and wherever there is information, there is experience." (Chalmers, 1996: 297) And a bit later:

> "It suggests a view in which the experiences in simple systems such as atoms are fundamental, and in which complex experiences are somehow the sum of such simpler experiences. While this is one way things could go, there is no reason that things have to go this way: complex experiences may be more autonomous than this suggests. In particular, the informational view suggests a picture on which complex experiences are determined more holistically than this" (Chalmers, 1996: 299).

In panpsychist terms, this is close to a position like panqualityism[55], which tends to deflate the subject of experience, reducing it to just structures of qualities.[56]

According to Chalmers, three different aspects of phenomenal states yield three different combination problems: the *subject combination* problem, the *quality combination* problem, and the *structure combination* problem (Chalmers, 2015).[57] The hardest of all problems seems to be the subject combination problem or *subject-summing*. If panpsychism is merged with IIT (or if IIT is interpreted as a form of panpsychism), perhaps IIT could help panpsychism solve some of these combination problems. Though it may seem possible to imagine IIT attributing to the understanding of how macro-qualities arise from micro-qualities

[55] Coleman's version of panpsychism in which the basic, intrinsic properties are qualities as 'unexperienced qualia'.

[56] Similarly, Cerullo (Cerullo, 2015) has pointed out that IIT's panpsychism is a version of panexperientialism. Cerullo argues that "theories of panexperientialism, therefore, measure proto-consciousness (or proto-mentality) rather than consciousness" (Cerullo, 2015: 8). So, in his interpretation, IIT would be a theory of "partial-panexperientialism". As a panpsychist, Rosenberg (Rosenberg, 2004: Chapter 5) expounds a close form of panexperientialism, the view that experience is everywhere in nature, but it is only in a subject of experience bounded in qualitative field. The boundaries of the field individuate subjects.

[57] Chalmers finds two additional problems: "There is the unity problem: how do microexperiences come together to yield a unified consciousness? There is the boundary problem (Rosenberg, 1998): how do microexperiences come together to yield a bounded consciousness?" Roelofs (Roelofs, 2015: 132) defines it as "Essential Boundedness (EB): The set of experiences belonging to any subject is bounded."

and how experience is integrated, at least *prima facie*, it is not so easy to see how the same could hold for subject-summing. Subjects of experience don't seem to be the entities that are open to combination or processes of fusion/merging. This could even be a conceptual impossibility. Of all three problems, I will be concerned with the third, that of the combination of subjects, since it is directly related to the issues of subjectivity.

One possible understanding of IIT would be as a form of *emergent panpsychism*. Chalmers points to the *emergentist* inclinations of IIT:

> Giulio Tononi's integrated information theory (2008), which puts forward a principle connecting degrees of integrated information with states of consciousness, can also be construed as a form of emergent panpsychism. If we see Tononi's principle as a fundamental law of nature, then it appears that macroexperiences are strongly emergent from certain physical configurations (Chalmers, 2016: 16).

Luke Roelofs views Tononi as a *physicalist-panpsychist*. He points out that a physicalist combinationist like Tononi would claim that information integration is the experiential bonding[58] but that we learn of it by observing brain structure, so this will leave primitivists about consciousness unsatisfied. For Roelofs, Tononi provides a theoretical rationale for boundedness of consciousness with *Exclusion* (Roelofs, 2015: 132), and this is an argument against combinationism.[59] Tononi defends *Exclusion* from phenomenology and parsimony. On the other hand, Shani interprets these authors (Koch, Tononi) as saying that information integration generates consciousness, which would be a "non-panpsychist idea" (Shani, 2015). Panpsychism assumes that sentience is combined and amassed into larger unities (macro-level consciousness) from smaller ones (micro-level consciousness), not from something non-sentient. Contrary to Shani's claim, it seems to me that IIT defenders take information to be imbued

[58] Phenomenal bonding relation.

[59] The position of constitutive panpsychists who defend the plausibility of the combination of experiences and subjects.

with sentience on all levels (or they should if they are to avoid the danger of radical emergence, the emergence of experience from something non-experiential).

There is no aggregation in IIT (Tononi & Koch, 2015). Integration axiom implies emergence in conscious experience. Subjectivity could be what unifies experiences from the inside. Also, one way to interpret *Exclusion* is as stating causal emergence in Integrated Information. I think it would be better to read *Exclusion* as stating that experience comes in discrete, unified subjects (individuals). If this integration is missing, then there is no subjectivity in IIT. It all depends on how we view subjectivity, as just relations between experiences (deflationary), or something more. In any case, on the panpsychist interpretation (with a panpsychist "twist"), IIT is better equipped metaphysically to deal with the hard problem of consciousness. If understood as a form of panpsychism, IIT avoids postulating radical emergence. These are the good sides of panpsychism that IIT inherits. This is all well, but what about subjectivity and subject-summing? Can the panpsychist IIT avoid these problems?

There are many ways one may try to solve the aforementioned combination problems. One common strategy used to answer the subject-summing problem is to *deflate* the subjects of experience, meaning that one denies that experiences need to have subjects "or at least denying that subjects are metaphysically and conceptually simple entities" (Chalmers, 2015: 271). If subjects are simple entities, fusion of such subjects into larger ones seems not so plausible.[60] To avoid these problems of subject-summing, I offered a solution based on subjectivity emergence in the previous chapter (Section 4.7.4).

[60] For discussion, see Coleman's (Coleman, 2013) anti-summing arguments. Some philosophers would argue that experientiality (the quality of experience, as a broader term, as opposed to subjectivity or the "subjective character" of experience) is the constant in consciousness. Even the weaker notion of the subject is abandoned in some panpsychist accounts. This is not so hard to understand, given that experience is something that we can more easily imagine being subject to processes like fusion and fission, merging and combination. Galen Strawson's theory of SESMETs is not very clear on this. Although he argues for the identity of subjects and experience, and even though persistent subjects are deflated to "thin subjects", they are not eliminated from experience. Subjectivity (and with it the "thin subject") is still an aspect of an experience that has to be reckoned with.

Therefore, in order to alleviate the pressure of subject-summing, panpsychists deflate the subject instead. The emergence of macro-subjects from micro-subjects is less plausible if subjects are fundamental, simple entities. If IIT is understood in panpsychist terms, we see the same worry expressed in the *Exclusion* postulate of Information Integration. Since IIT accommodates split-brain cases and fission of subjects, it would then seem to subscribe to the fusionist panpsychist accounts akin to that of one William Seager (Seager, 2010). Some, like Roelofs and Mørch, suggest that the phenomenal bonding relation[61] could be Information Integration. So, IIT could help explain the fusion view (Seager, 2010; Mørch, 2014). Macrophenomenal properties are fusions of microphenomenal ones. Micro-subjects fuse into macro-subjects at a later time (diachronic emergent combination).

It seems arguable that Integrated Information Theory would subscribe to a deflationary account of subjectivity. It would reduce subjects to experience and its relations, at best, even given its panpsychist interpretation. In the next section, I will show that IIT lacks this important aspect that every theory of consciousness should have, that is, being able to account for the subjective character of consciousness. Panpsychist theories that are supposed to help IIT get off the ground are also plagued by the same problem. In this Chapter, I will confine my criticism to Integrated Information Theory and its phenomenological axioms.[62]

5.4 INTRINSIC PERSPECTIVE?

The biggest drawback of IIT and a persistent problem for the theory, even if it is improved by panpsychism, would still be that it posits no subjectivity, no point-of-view-edness. This is the claim on which the present chapter has been based. It would seem a conceptual truth[63] that every experience has an

[61] Phenomenal bonding relation, a special kind of relation that holds between subjects of experience (Goff, 2009), or between experiences within a subject (this could be the "co-consciousness" relation).

[62] I criticize versions of panpsychist deflation of subjects in Nešić, 2017.

[63] See Chalmers, 2015 for discussion.

experiencer (subject). Any good theory of consciousness needs to explain subjectivity and say something about the nature of subjects. Even if one is not a defender of subjects as fundamental ontological entities, one has to acknowledge that something broadly construed as *"subjective character" of consciousness* is very likely to exist. So, "any satisfying theory of consciousness has to account for the first-personal access to our own consciousness" (Zahavi, 2005: 13). The hard problem of consciousness also demands this.

What is missing in IIT, and what could help us make sense of the boundedness of experience and, ultimately, make sense of Integrated Information, is the subjective character of experience. Subject-summing would be left untouched even if IIT could help panpsychism solve the structure combination and quality combination problems. Adherents of IIT, at least, seem to be aware of the need for this subjective aspect. "Information — the ability to discriminate among many alternatives — may thus be essential for consciousness. However, information always implies a point of view, and we need to be careful about what that point of view might be" (Tononi, 2008: 218). In IIT the maximally irreducible conceptual structure specified by a complex exists intrinsically (from its own *intrinsic perspective*). Shanahan also points out that IIT: "is unable to account for the sort of self-knowledge it takes to be axiomatic" (Shanahan, 2015: 9).

Maximally irreducible conceptual structure is identical to the integrated experience in the same way the subject is identical to its stream of consciousness, to the total experiential field. Every experience, and thus every piece of information, is always a part of some subject, always marked by first-person givenness, so every piece of information in a system is also marked by its subjectivity and there can be no other conscious subsystem (*Exclusion*).

This could mean that subjectivity is an aspect *only* of Integrated Information and of the system that realizes it. There is the identity of phenomenological properties of experience and causal/informational properties, so it would seem that intrinsic perspective is not present in every experience as such but *in the integrated whole* of information/experience. Authors state

that maximally irreducible conceptual structure is identical to its experience (Oizumi *et al.,* 2014: 3). If Integrated Information is not regarded as something more than the combination of experience/information, IIT is in danger of becoming a deflationist theory of consciousness. It would reduce subjects to experience or completely remove them from the metaphysical picture.

This brings us to the highly debated question of *unity of consciousness* and what it is. How should we understand the unity of consciousness? Chalmers and Bayne (Chalmers & Bayne, 2003) define it with *The Unity Thesis*: "Necessarily, any set of conscious states of a subject at a time is unified". If consciousness is unified, one needs to explain why there is such a striking unity, and what brings all the conscious states into one unified experience. One way to explain the unity of consciousness is with the "co-consciousness" relation. One proposed solution to the subject-summing problem (that could be endorsed by proponents of IIT) is the *phenomenal bonding* strategy (Goff, 2009), positing a special kind of relation holding between subjects, though the nature of this relation is unknown to us. This is because we can only introspect within a subject. Chalmers has proposed that the phenomenal bonding could be the "co-consciousness" relation. But what this "co-consciousness" relation really is, and how do we understand it? And how is it different from phenomenal bonding? One could ask if it is helpful in explaining the unity of consciousness, or if it just states a fact.

IIT acknowledges the existence and importance of unity of consciousness, as this is stated in the second and the fourth axiom. The basic claim in IIT is that consciousness (and information) is integrated, but there is no explanation as to why that is the case. IIT could be compatible, on the ground of *Exclusion*, with a subjectivist account of consciousness. But, the subjective character that is arguably present in our experience and is not mentioned by the current version of IIT is the first-person givenness or *mineness of experience* (Zahavi, 2005).[64] The authors of IIT claim that Information Integration is the relation that keeps the subject

[64] Guillot argues that what Zahavi has in mind when talking about mineness is actually for-me-ness.

"bounded". They postulate the identity of consciousness and Information Integration.[65] If we say that the subject is Integrated Information, this does not include any mineness and would be in line with a deflationist theory of the self/subject.

On the phenomenological understanding, subjectivity is like a dimension or a space that grounds any phenomenal relations and experiential properties. Subjectivity could be the space where experience is manifested, and the relation of the subject to its experiences could be the same as that of spacetime to its material objects.

Saying that subjectivity is the same as Information Integration does not reveal the full truth about subjectivity. Information Integration, like phenomenal bonding, is a *relation*. If there is something like pre-reflective self-awareness in the sense that philosophers like Zahavi, Nida-Rumelin, Siewert, Guillot and others understand it, at least some routes of defending subject-summing seem to be blocked. One way to imagine how subject-summing is possible is by way of the "co-consciousness" relation (Dainton, 2011). Contrary to that, I think that the phenomenology of self-awareness tells us that the fact about "co-consciousness" is based on the fact that experiences belong to the same subject (not the other way round). This shows that phenomenal bonding as "co-consciousness" is implausible.

Slors and Jongepier (Slors & Jongepier, 2014) argue that the mineness of experience is a product of the *external structure of experience*. They defend a reductionist, *coherentist* account of mineness and argue against Zahavi's account of mineness. Due to their approach, these philosophers deny there is any phenomenological datum to mineness. There are no minimal selves in their theory. Still, they would agree with Zahavi that "thoughts are endowed with a first-personal givenness — how can they fail to be?" (Zahavi, 2014: 216). From what we have seen, this seems to be the case in Integrated Information as well, and a defender of Integrated Information Theory could endorse the coherentist account.

[65] One could imagine a possible (and rather generous) interpretation of IIT that would state that Φmax measures the value of subjective character, the level of subjectivity and, hence, consciousness of a system.

Similarly to the situation in Coleman's panqualityism, it is conceivable that a system with Integrated Information lacks awareness — though it has information/experience integration, it would be an "awareness zombie" (Chalmers) and not a subject of experience.[66] Attempts to "functionalize" awareness, and this would be the case in IIT, also eradicate the phenomenology of awareness, which is very unpalatable.[67]

More work is to be done on future versions of IIT, in order to properly fit the subjective character of consciousness into the architecture of Integrated Information Theory.[68] There may be a more fundamental problem of Integrated Information Theory if one is to understand it as a panpsychist theory, and I will tackle this issue in the following chapter.

[66] See Coleman's paper in Bruntrup & Jaskolla, 2016 for discussion.

[67] Therefore, IIT seems to be threatened by the *nonsubject/subject gap*. See Chalmers, 2016 for a discussion on these issues.

[68] I posed this problem for IIT in Nešić, 2018a.

6. SELVES AS STRUCTURES

> Consciousness gradually loses its coherence. One's center gives way. The center cannot hold. The "me" becomes a haze, and the solid center from which one experiences reality breaks up like a bad radio signal. There is no longer a sturdy vantage point from which to look out, take things in, assess what's happening. No core holds things together, providing the lens through which to see the world, to make judgments and comprehend risk. Random moments of time follow one another. Sights, sounds, thoughts, and feelings don't go together. No organizing principle takes successive moments in time and puts them together in a coherent way from which sense can be made. And it's all taking place in slow motion.
>
> *The center cannot hold: my journey through madness*
> (Saks, 2007: 13)

6.1 ONTIC STRUCTURAL REALISM

A new metaphysical theory of consciousness and selfhood (Beni, 2019) that bears the mark of structuralism is on the rise. The usual ontological suspects picked out by most of the philosophers of mind when it comes to the question of the nature of the self are the substance (substantivalism) and the bundle. There are also those who claim that there is no such thing as the self (Metzinger, 2003) in any substantival sense, or that it is an illusion (Frankish, 2016). Building on a rather familiar landscape of self(-less) theoretical possibilities, Beni defends a realist theory of selfhood as an informationally regimented structure. He expounds an information-theoretic structural realist version of both the theory of consciousness and the self. One aspect of this structuralist account denies intrinsic properties; the other denies objects.

The focal point of this Chapter is pinpointing the challenges inherent in the mentioned position and address these challenges while adhering to the principles of structuralism. The overall structure of the chapter is as follows: In Section 6.1, I will elucidate the foundations of the structuralist theory, specifically the framework of ontic structural realism. Moving to Section 6.2, I will unpack Beni's structural realist theory concerning the self and consciousness. Section 6.3 will delve into the analysis of his preferred metaphysics of consciousness, highlighting a predicament for Beni's panpsychist interpretation. In Section 6.4, I will introduce Mindt's (Mindt, 2019, 2021) amendments to the Integrated Information Theory, incorporating intrinsic structure, and propose how this can be applied to enhance Beni's theory, enabling it to offer a more comprehensive and satisfying resolution to the hard problem of consciousness. Section 6.5 will delve into the discussion of neutral monist ontology, exploring how the information-theoretic structuralist theory could be refined through integration with neutral monism.

Beni's theory of consciousness/selfhood is founded on the framework of *Ontic Structural Realism (OSR)*. Structural realism is a successful theory of philosophy of science which posits that the structures described by scientific theories are real.[69] The emergence of this view was motivated by problems of pessimistic meta-induction and metaphysical underdetermination.[70] Structural realism comes in several flavours. *Epistemic Structural Realism* (ESR)[71] is the claim that all that we *know* is structure (there may

[69] See Ladyman, 1998 and Worrall, 1989. For an insight into recent debates on scientific progress and different forms of scientific realism see Dellsén, 2018, Niiniluoto 2019, Bird, 2016.

[70] Ontic SR parts ways with traditional substantivalism and object-oriented metaphysics. SR emerged as an answer to pessimistic meta-induction and metaphysical underdetermination in physics. Beni (Beni, 2019: 44) points out that the object-oriented version of scientific realism could not resolve these problems satisfactorily, but SR could. His claim is that a similar problem of the metaphysical underdetermination occurs in psychology pertaining to the nature of the self. Beni's motivation for adopting SR is that it could resolve this problem concerning the self, as well. More on the problem of metaphysical underdetermination of the self in Footnote 10. I thank one of the reviewers of this journal for pressing me to say more about this point.

[71] The following is a garden-variety distinction made between different versions of structural realism. See Ladyman, 2020.

be hidden entities realizing the structure, but we know not of their nature). The *Ontic Structural Realism* (OSR) is the claim that reality is fundamentally relational or structural — there are no objects or, if there are any, they are grounded in the structure (Ladyman et al., 2007; French, 2014). OSR can be of the eliminativist and non-eliminativist kind. *Eliminativist Structural Realism* goes the furthest in its gung-ho charge on traditional metaphysics by professing: "All there is, is structure!"

Moderate or *non-eliminativist ontic structural realism* (Esfeld & Lam, 2008) is not so extreme. Those who endorse it, like Esfeld, maintain that objects are to be characterized only by the relations in which they stand. In this version of OSR, there are both objects and relations, so it respects that relations require relata, but objects have no fundamental intrinsic properties[72] — objects only bear relations and relations account for identity conditions instead of intrinsic properties.[73]

As opposed to OSR, in traditional metaphysics, individual objects were considered substances and their individuality was understood as intrinsic and primitive, independent of the external. To individuate objects, one could use an intrinsic property, haecceity (primitive thisness, individual essence) or a bare substratum.[74]

[72] There are relational/extrinsic (structural) and nonrelational/intrinsic properties (and these can be categorical or fundamentally dispositional) (e.g. Seager, 2006). According to Jeagwon Kim (Kim, 1982), intrinsic property is to be understood as the property that belongs to an object that does not coexist with any contingent object distinct from itself (lonely or unaccompanied object). Lewis defines an intrinsic property as property "which things have in virtue of the way they themselves are". Things have an extrinsic property "in virtue of their relations or lack of relations to other things" (Lewis, 1986: 61). In Lewis and Langton (Lewis & Langton, 1998), it is argued that "intrinsic" properties are logically independent of both loneliness and accompaniment. Francescotti (Francescotti, 1999: 608) considers that F is an intrinsic property $=_{df}$ necessarily, for any item x, if x has F, then there are internal properties $I_1,...,I_n$ had by x, such that x's having F consists in x's having $I_1,...,I_n$.".

[73] The master argument for intrinsic properties given by Esfeld and Lam goes as: "(1) Relations require relata, that is, objects that stand in the relations. (2) These objects have to be something in themselves, that is, they necessarily have some intrinsic properties over and above the relations that they bear to one another — even if the relations do not supervene on the intrinsic properties and even if we cannot know the intrinsic properties" (Esfeld & Lam, 2008: 29).

[74] Speaking of consciousness and subjects of experience, there are many positions, including panpsychism, that are serious about consciousness and

Esfeld and Deckert (Esfeld & Deckert, 2018) defend moderate OSR and assert that it is a "misconception" to understand ontic structural realism as a position that goes against object-oriented metaphysics. Ontic structural realism is a stance that goes against the *property-oriented metaphysics* that has been dominant in philosophy since Aristotle. There really is "no need to admit physical properties at all". There are objects in a thin sense, "standing in the relations is all there is to these objects — the relations are their essence (cf. the moderate ontic structural realism set out in Esfeld (Esfeld, 2004), Esfeld and Lam (Esfeld & Lam, 2008, 2011))" (Esfeld & Deckert, 2018: 7). Objects understood in this are not objects in the proper sense; they possess no "thisness" and there is no *bare substratum*.[75]

I go to length to explain these forms of OSR both for introductory purposes and to get a better grip on how to situate Beni's account. His theory comes close to the *non-eliminativist* OSR, which he acknowledges (Beni, 2019: 63). More about his non-eliminativist version of OSR will be said in Section 3. What is important to note here is the following: from everything said, it is clear that all these forms of Ontic SR dispense with intrinsic properties. I highlight this fact because it will be crucial for my arguments in the coming sections (especially in Sect. 4).

Now, some philosophers (Ladyman et al., 2007) have pushed towards developing structuralist theories in domains other than physics: chemistry, biology, economics, and cognitive science. Beni's *Structural Realist theory of the Self* (SRS) *and consciousness* is part of this tendency.

subjectivity but do not posit subjects as enduring substances, although they still have intrinsic properties. See Seager 2006, Dainton 2008, Mørch 2018.

[75] Problems for structural realism could arise from quiddities. Ungraspable quiddities would undermine *ontological* structural realism, while graspable quiddities would undermine both *ontological* and *conceptual* structural realism. Now, quiddities need not undermine *epistemological* structural realism, as long as the distribution of the quiddities is not knowable (Chalmers, 2012: 422).

6.2 STRUCTURAL REALIST THEORY OF THE SELF

Traditional views on the metaphysics of the self were *the substance view* and *the bundle view*. Following the Phenomenological tradition, an additional view has been postulated — the *minimal self view*. Now, Beni proposes a radically different *metaphysical* theory of the self, based on the ontology of *structures*.[76]

To revisit the classic distinction in understanding the self, we initially separated it into the bundle and substance perspectives. To this, I introduced the *phenomenal self* derived from the Phenomenological tradition. In the bundle view, the self is either identical to or constituted by bundles of experiences. The individuation of a subject is determined by its experiences, with the identity of a subject defined by the relationships between mental states. Selves are essentially collections of properties, as outlined by Dainton. If the self is considered a "minimal self," it is essentially synonymous with the subjectivity of experience. Zahavi advocates for experiential minimalism, asserting that the for-me-ness or first-personal giveness aspect of phenomenal consciousness represents the minimal experiential self.

An *individual substance* is a self-individuating entity that metaphysically unifies and individuates properties as their bearer. If the experiencing subject is an individual substance, then it is a kind of metaphysical entity that acts as a bearer of experiential properties, upon which experiential properties are instantiated, and that is not itself a property. Hence, the instantiations of experiential properties in subjects are types of events, namely experiences (e.g. Lowe, 1996; Nida-Rümelin, 2017).

Beni argues against the bundle and substance views, but it is also his intention to go beyond what he calls *eliminativism* and *pluralism* that are evident in the contemporary understanding of the nature of the self. Metzinger's eliminativism propounds there is no self, only a "self-model", a mere phenomenal image produced by neural representations, while Gallagher's pluralist *pattern theory* (Gallagher, 2013), in which there are "multiple

[76] Beni's view is radically different in the same way that OSR is radically different from traditional object-oriented ontology.

co-existing and loosely related self-patterns", does not account for the dynamical relations between various self-patterns. Beni contends that eliminativism and pattern theory do not provide a well-posed metaphysical alternative to substantivalism. All of these theories pull in different directions to the effect that they distract from a unifying ontological account of the self.

The structural realist theory of the self does not endorse full-fledged realism concerning all aspects of the self. Beni asserts that his intention is to defend a modest version of realism about the basic structure of the self that can be specified in terms of embodied informational structures, or structures realised by mechanisms of information processing in the brain and the environment. There is a diversity of scientific accounts of the self, and Beni has been the first to point out that there is a *metaphysical underdetermination of the self*. This underdetermination is what motivates him to seek out a structuralist theory of the self.[77] So, this is to be a philosophical theory of selfhood based on what our best theories of contemporary cognitive sciences tell us about the self.

Beni introduces his view as a *non-eliminativist* version of OSR that gives priority to basic structures, which means he does not want to dispense with objects altogether, but to retain an ontologically thin notion of individual objects (Beni, 2019: 63). What remains are "weakly discernible individual selves" where thin notions of "non-structural features of the self can be identified in virtue of the specific location that they occupy in the infrastructure of the selfhood" (Beni, 2019: 123).[78]

[77] The problem of metaphysical underdetermination is encountered in the field of the philosophy of physics. The metaphysical underdetermination comes from quantum statistics cases which result in incompatible metaphysical consequences (French, 2019). A solution for this problem is to show that there is a common structure that underpins the individualistic and the non-individualistic notion of objects. This is where the Ontic SR comes into play (Beni, 2019: 110). Beni finds similarities between the state of underdetermination in physics and in the neuroscientific accounts of the self. Therefore, an analogous threat pertains to the philosophy of self (a form of metaphysical underdetermination "breaks out" in this field anew). This underdetermination is caused by the heterogeneity of neuroscientific theories of consciousness, like the Integrated Information Theory, the resting-state-based theory, and the FEP-based theory of consciousness. For a detailed exposition of this problem, see Beni (Beni, 2019: Sect. 2.2.2 and 3.8).

[78] The sense of agency, the sense of ownership, and mineness.

Beni admits some non-structural elements in his theory. What does it mean to have such elements in the framework?[79] They still get their identity from their location in the structure:

> According to SRS, the self is the infrastructure that subsumes various self-patterns. Self-patterns or various aspects of the self are featuring in the infrastructure of the self. Aspects and features of the self (i.e., self-patterns) could be identified mainly by virtue of their location within the infrastructure of the self. (Beni, 2019: 127)

Beni subscribes to the *information-theoretic structural realism* (ITSR) (Ladyman et al., 2007). According to these authors, structures are to be characterised by *real patterns* (Dennett, 1989: 38–42; Ladyman et al., 2007: 252; Mindt, 2019: 120–104). Real patterns were defined by Dennett (Dennett, 1991, 1989) as patterns that are objective and exist out there, to be detected by our observations of the natural world. However, there is always an intentional stance to these observables. Ladyman and colleagues claim that real patterns are "the last word in ontology, and there is nothing more to the existence of a structure than what it takes for it to be a real pattern" (Ladyman et al., 2007: 178). Some real patterns behave like objects, and some behave like events and processes.

So, when it comes to the structural realist theory of the self/consciousness, Beni finds it more natural to specify the underlying structures of the self in information-theoretic terms. SRS is a specific form of informational structural realism that grounds the informational structures in a cognitive system that can be coupled with the environment (echoes of enactivism and active inference).

In addition, he tries to employ various accounts of consciousness, e.g. the Integrated Information Theory of conscious-

[79] One could envisage an argument against structuralism claiming that selves are the only real objects we know about. Nida-Rümelin (Nida-Rümelin, 2017), for example, would argue that we have pre-reflective self-awareness, an awareness of oneself as an experiencing subject — we are aware of ourselves as unifying simultaneous and subsequent experiences. She defends the view that the self is an unchanging experiencing subject, a special kind of substance that has a non-descriptive individual nature. See Sections 1.2 and 2.3 of this book.

ness (Tononi et al., 2016),[80] resting-state-based theory (Northoff, 2018), and free-energy-based theory (FEP, Friston, 2010). He goes on to show how they could be unified by invoking a structural realist strategy. Beni argues that to overcome the problem of metaphysical underdetermination of consciousness, we should find the *underpinning structure* of consciousness. Beni thus offers a structural realist account of the phenomenal aspects of the self and consciousness, even intentionality.

6.3 METAPHYSICS OF CONSCIOUSNESS

The adequacy and plausibility of a theory of consciousness become especially problematic when one tries to explain phenomenality in structuralist terms. This is the reason why I think that the dreaded and, perhaps, intractable *hard problem of consciousness* opens up for Beni. Chalmers argues that reductive explanations concern only structure and function but do not explain the conscious experience that accompanies it (Chalmers, 2010: Chapter 1).

Beni does not pose the hard problem of consciousness explicitly in his book.[81] Let us investigate some more the metaphysics behind his theory of consciousness. On one occasion he confesses that he is "advocating a limited form of panpsychism" (Beni, 2019: 197), a kind of *pancomputationalism-cum-panpsychism*,[82] where all information processing (to some extent)

[80] Mindt argues that IIT is unable to properly answer the hard problem of consciousness in its present form. It falls victim to the structure and dynamics argument (Mindt, 2017).

[81] In his account, Beni employs IIT and free energy principle as structural theories of consciousness. If free-energy-based theory is also conceived as structuralist it would be hard to see how it is panpsychist, and Beni does claim that Friston's theory supports panpsychism (Beni, 2019: 183–4). In more recent work, Beni (Beni, 2021a) posed the hard problem of consciousness for FEP and pushed for the critique of the so-called *Markovian monism*, in that it is not a viable metaphysical theory of consciousness or a good answer to the hard problem. In another paper, he uses the scientific literature around the Free Energy Principle to reconstruct two well-known arguments for panpsychism — the argument from continuity and the argument from intrinsic nature (Beni, 2021b).

[82] Perhaps, what Beni means by "panpsychism" (and "pancomputationalism-cum-panpsychism") is the following: where there is information processing,

involves phenomenal experience. He claims that this view seems to follow from his information-theoretic account of the structure of the self. His panpsychism is also in line with IIT, Beni contends.[83] This view of Beni's, although important for his theory, would need some more unpacking. Unfortunately, he does not dwell much on it, and his panpsychist view is not explicated in detail. What seems to be problematic is that Beni's is an ontic structuralist theory, and this does not form a fertile ground for a panpsychist ontology.

Panpsychism is the view that phenomenality/consciousness is fundamental and pervades the physical universe. Everything has a modicum of consciousness or properties that are similar to conscious (proto-conscious). Modern panpsychism comes as a promising position that could overcome the deficiencies and problems of both physicalism and dualism, and that could constitute a more plausible answer to the hard problem of consciousness (e.g. Strawson, 2008; Mørch, 2014; Roelofs, 2014; Chalmers, 2015; Alter & Nagasawa, 2015; Bruntrup & Jaskolla, 2016; Goff, 2017; Seager, 2019).

Can there be a structuralist version of panpsychism? How could panpsychism and *Ontic Structural Realism* (even of the moderate kind) be compatible? Ontic SR is a form of *non-categoricalism* (Ladyman et al., 2007). Panpsychists usually argue against OSR, and for categoricalism (cf. Mørch, 2018; Brüntrup, 2011; Seager, 2006)[84]. Panpsychism has the requirement of *intrinsicality* and structural realism does not — even the moderate version

there is phenomenal experience/consciousness. Whenever there is life, there is conscious experience. In that case, Beni would subscribe to the *mind-life continuity* thesis. There are some clues to this in Beni (Beni, 2021b).

[83] Orthodox interpretation of IIT is panpsychist, but since IIT was formulated this interpretation has been highly controversial. McQueen (McQueen, 2019) argues that orthodox interpretations of its own ontological and epistemological basis should be rejected for an interpretation-neutral formulation. Cf. Mørch 2018.

[84] This is how Mørch defines OSR: "Ontic structural realism is the view that all physical properties are purely structural or relational, but that relations do not need relata with non-relational properties; rather, physical relations can subsist on their own, or at least prior to their relata such that the relata are constituted by their position in a relational structure and would have no reality outside of it" (Mørch, 2018: 4). The alternative to OSR (in which structures are fundamental) is *dispositionalism*, where dispositions or powers are fundamental, non-reducible to structural or categorical properties. Dispositionalism can answer the

of structuralism (Esfeld & Lam, 2008) clearly eliminates intrinsic properties.

In the literature, intrinsic and categorical are used almost synonymously, (intrinsic and categorical nature). Still, these are not the same. Intrinsic properties are properties that are constitutively independent of the properties of other things and categorical properties are understood as those independent of other properties, including "circumstances and manifestations" (Mørch, 2018: 5).[85]

Categoricalism can be viewed as the position that dispositions require categorical grounds or realizers. Categorical properties are also seen as non-dispositional properties. Dispositionalism is compatible with intrinsicalism because powers can also be intrinsic — irreducible powers can exist unmanifested (Molnar, 2008).[86]

Perhaps, panpsychism should be avoided because it is admitting too many entities as sentient. To his credit, Beni does expound a "limited form of panpsychism". Integrated Information theory had the corollary that even very simple objects, like photodiodes, are conscious and this was the source of criticism towards

question what distinguishes physical from mathematical structure — physical structure is realized by powers.

[85] I discussed what intrinsic properties are in a previous footnote. Russell (Russell, 1927) held that physical events have an intrinsic character beyond the structuralism of physics. In structuralist views of physics, which eliminate intrinsic properties, "all the things in the world will merely be each others' washing" (Russell, 1927: 325). Recently, many views (inspired by Russell's positions) under the umbrella term *Russellian monism* have appeared claiming that "matter has intrinsic properties that both constitute consciousness and serve as categorical bases for the dispositional properties described in physics" (Alter & Nagasawa, 2015: 1). Chalmers' type-F monism is also a form of Russellian Monism or panprotopsychism (Chalmers, 2010).

[86] The same assertion can be found in Esfeld and Deckert (Esfeld & Deckert, 2018: 54) about dispositions as intrinsic properties and that is why they are not admitted in structural realism. It is also the reason why Esfeld and Deckert say that not even physical properties need to be admitted, just relations. Endorsing dispositionalism goes beyond structuralism, amounts to something more than ontic structural realism. Chakravartty goes through all the ontological alternatives of object ontologies, between substance realism and eliminativism (thick and thin): substances, bundles, dispositions (*dispositional essentialism*: there is a generally *intrinsic* potential for relations, causal powers investigated by the sciences are generally intrinsic properties). He defends *semi-realism* (Chakravartty, 2007) incorporating the *bundle view* through a dispositionalist account (French, 2014: Chapter 7 entertains the possibility of "bringing back the bundle").

the panpsychists interpreting IIT.[87] However, it could be that this is a problem just for IIT, though not for all of panpsychism. On the other hand, IIT could give a more precise answer to the question how much consciousness there is in a physical system.[88]

Given that panpsychism is bound with intrinsicality, Beni's endorsement of panpsychism is incompatible with his commitment to Ontic SR. Epistemic SR would not be a problem for a panpsychist, since panpsychism is comfortably compatible with *Epistemic Structural Realism*. Still, adopting Epistemic SR is not available to Beni, given his explicit commitments to Ontic SR.[89] In order to have a full-fledged panpsychist theory, Beni would need to give up his allegiance to OSR, something that is at the very core of his structuralist theory of the self and consciousness. I do not think Beni is ready to make this move and I will argue that a different metaphysical theory of consciousness is a better match for Beni's structuralism.

What alternatives does Beni have, aside from embracing panpsychism? A conflict arises between Ontic Structural Realism (Ontic SR) and panpsychism due to the concept of intrinsicality. If we reject intrinsicality, as ontic structuralists do, our choices for compatible metaphysical theories of consciousness become limited. Can this structuralist theory be seen as a form of physicalism? The challenging issue for understanding consciousness, known as the hard problem, surfaces once more. It's crucial to note that this hinges on our understanding of what qualifies as the physical and what physicalism entails.

While physicalism appears to be a prevalent metaphysical theory of consciousness, defining it poses a challenge.[90] What renders a process or entity as physical? One could adopt the widely recognized definition by Stoljar, employed by both Chalmers

[87] Is every living creature also conscious? What creatures are endowed or imbued with consciousness? Consider Godfrey-Smith (Godfrey-Smith, 2016) for discussion. Mindt (Mindt, 2021: 9–10) is cautious when discussing the relationship between life and consciousness.

[88] For example, what is the difference in levels of consciousness in a newborn baby, an animal like a bat, or in vegetative patients with just "islands" of brain activity?

[89] Goff (Goff, 2021) is explicit about his commitment to Epistemic SR, for example.

[90] Physicalism and materialism are usually taken to be synonymous.

(Chalmers, 2003) and Mindt (Mindt, 2021). This *theory-based conception of the physical* posits that physical entities (processes, properties) are those recognized by a physical theory.[91] Mindt explores Lewis' (Lewis, 1983) definition of physicalism as the metaphysical view that there is "nothing over and above the physical" (Lewis, 2021: 2). When we refer to the physical, we are essentially describing structural and dynamical processes/properties. Chalmers (Chalmers, 2010: 120) explains that this involves depicting the world in terms of its underlying spatiotemporal and formal structure and dynamic evolution over this structure, extending beyond mere properties of spatiotemporal entities to encompass dynamical, causal features associated with third-person observation. Mindt (Mindt, 2021: 3) suggests labelling these as *extrinsic features* of a system.

I find the hard problem to be a significant challenge for physicalism. This brings us to Mindt's proposals for addressing the hard problem through *intrinsic structure*.

6.4 INTRINSIC STRUCTURE

Can Beni address the hard problem of consciousness while remaining firmly rooted in structuralism? I propose that a potential solution to this challenge can be found by examining how Mindt (Mindt, 2021) has defended Integrated Information Theory against the structure and dynamics argument. Similar to IIT, Beni's theory operates within the framework of information theory, offering an information-theoretic account of consciousness. If Beni's theory relies solely on a structural (and dynamical) understanding of information, it might seem ill-equipped to tackle the hard problem.

Mindt points out that the heart of the hard problem of consciousness lies in the *structure and dynamics* (S&D) argument presented by Chalmers (Chalmers, 1996, 2003). Chalmers contends that physical truths alone cannot fully explain

[91] The theory-based conception: "A property is physical iff it is the sort of property that physical theory tells us about" (Stoljar, 2017).

consciousness because they focus only on structure and dynamics (Chalmers, 2003; Alter, 2016), presenting this as an argument against physicalism/materialism.[92] Physical descriptions characterize the world through analysable terms such as formal, spatiotemporal, logical, and mathematical, as well as nomological aspects like laws and causation, all of which pertain to structural and dynamical processes/properties. "Structure and dynamics" in this context refer to "spatiotemporal and formal structure, and dynamic evolution over this structure" (Chalmers, 2003: 258), often considered a fitting description of the "physical" (following Russell). When a system possesses consciousness, there is a subjective experience associated with it.

Mindt suggests expanding our understanding of structure and dynamics by introducing another category of properties — *intrinsic structure and dynamics*. He argues for a more nuanced view of structure and dynamics than what Chalmers presupposes in the S&D argument. Mindt's aim is to dissolve the hard problem of consciousness by demonstrating that not all structure and dynamics are equal.

Mindt identifies different types of information relevant to his argument: *syntactic information* (aligned with Shannon's entropic notion of information), *semantic information*, and *intrinsic information* (Mindt, 2021: 4). A successful information-theoretic explanation of consciousness, one that overcomes the hard problem, should elucidate the relationships between these types of information. Mindt addresses the connection between syntactic and semantic information by adopting the complexity sciences' approach, which presents an evolutionary narrative of how meaning is generated. Kolchinsky and Wolpert (Kolchinsky & Wolpert, 2018) define semantic information as "the information that

[92] Argument's principal claims are: "First: physical descriptions of the world characterize the world in terms of structure and dynamics. Secondly: from truths about structure and dynamics, one can deduce only further truths about structure and dynamics. And thirdly: truths about consciousness are not truths about structure and dynamics." (Chalmers, 2003: 120). Alter (Alter, 2016: 2) defends Chalmers' S&D argument and formulates it as: "1. All physical truths are purely structural. 2. From purely structural truths, one can deduce only further purely structural truths. 3. Some truths about consciousness are not purely structural. 4. Therefore, there are truths about consciousness that cannot be deduced from (i.e., are not a priori entailed by) the complete physical truth."

a physical system has about its environment that is causally necessary for the system to maintain its own existence over time." They view information supporting the maintenance of a state of non-equilibrium as meaningful or semantic, aligning with the free-energy principle (Friston, 2010).

Mindt contributes to the discussion by proposing a perspective on how semantic information should be understood: "Those syntactic relationships which exist between an organism/system and its environment have value (i.e., have semantic content) if the syntactic information has the result of helping the system causally maintain its existence over time" (Mindt, 2021: 5). The emphasis is on identifying properties that play a causal role in sustaining a system's existence over an extended period. Mindt explicitly endorses a non-standard interpretation of semantics as the relationship between a system and its environment (Mindt, 2021: 7). This form of semantic information is not constrained to being propositional or epistemic; rather, it is substrate-neutral and doesn't adhere to a single standard. Mindt characterizes himself as a pluralist regarding various conceptions of semantic information, each capturing different levels of explanation. However, he contends that this type of information alone is insufficient for a comprehensive information-theoretic explanation of consciousness and experience since it primarily deals with meaning from an extrinsic or *external perspective*.[93]

Furthermore, there is a crucial progression from considering syntactic and semantic aspects to delving into the *intrinsic or internal perspective*. This involves moving beyond meaning that is "extrinsically interpreted from the outside" to understanding the meaning intrinsically within the system. Some systems may possess an internal viewpoint on their own processes, suggesting that there could be a subjective experience of the meaningful causal states that sustain the system's survival over time.

[93] He (Mindt, 2021: 14) differentiates between *extrinsic* and *intrinsic* structure and dynamics (S&D), and between *external* and *internal* (meaningful S&D properties), depending on the perspective one takes on the target system. From the external perspective, there are meaningful and meaningless S&D properties of a system. *Intrinsic S&D* can be interpreted from an external perspective, and this is connected to the semantic notion of information. There is a meaningless (non-meaningful) variant of external S&D or *extrinsic* S&D.

Mindt contends that Integrated Information Theory (IIT) can offer us the concept of intrinsic information and intrinsic structure and dynamics (S&D). Unlike other theories, IIT focuses on the subjective aspect of *something it is like* for a system to have meaning, delving into the intrinsic structural and dynamical features of a system from its internal perspective (such as my consciousness of the semantic features of myself or any conscious system's awareness of itself).

The crucial question posed by IIT is that of difference between the systems with intrinsic and extrinsic meaning. According to Mindt (Mindt, 2021: 13), the answer lies in the system's possessing the right intrinsic cause-effect power. Some systems, when examined externally, may only exhibit extrinsic structure and dynamics. These features are the subject of Chalmers' criticism, as they fail to account for the qualitative aspect of consciousness—the "what-it-is-likeness."[94]

Mindt's argument suggests that, given a specific understanding of structural and dynamical properties, Integrated Information Theory (IIT) falls short of resolving the hard problem of consciousness; it does not provide a satisfactory answer. However, Mindt posits that with a more nuanced comprehension of structure and dynamics, IIT stands a better chance of overcoming this problem. When two systems possess semantic information, the critical question arises: do both systems have experiences, and what distinguishes the one having an internal perspective from the other that does not? Mindt's proposed framework for structure and dynamics allows for a natural inquiry into these questions (Mindt, 2021: 15). It is essential to note that the success of Mindt's suggestions is contingent on their effectiveness. He does not assert that IIT is the optimal theory for capturing the properties of systems with internal structure and dynamics but maintains that it is heading in the right direction.

While Mindt acknowledges that detailed work is required to articulate how intrinsic structure and information contribute to explaining experience without relying on traditional intrinsic

[94] The extrinsic structural and dynamical properties only indicate a system's syntactical features and are the same features that Chalmers calls structural and dynamical in his characterization of physical explanations.

properties, he contends that adding an internal perspective to yield intrinsic information aligns with a naturalistic, information-theoretic approach to explaining consciousness. Despite potential reservations about whether this addition enriches the explanation of conscious experience sufficiently, both I and, I believe, Mindt argue that this proposition represents a promising step toward a naturalistic explanation and a better chance of tackling the hard problem.

Within the conventional understanding of structure and dynamics, as per Chalmers' argument, a structuralist theory like IIT seems incapable of addressing the hard problem. However, if one endorses a more nuanced depiction of structure and dynamics, such as Mindt's proposal, IIT and other structuralist theories may no longer be vulnerable to the structure and dynamics argument. Nevertheless, Mindt acknowledges that much more work is necessary, and the inclusion of these notions of structure and dynamics provides a potential rather than a definitive solution to the hard problem. Starting from IIT and the hard problem, Mindt contends that recognizing different types of "structure and dynamics" necessitates a departure from physicalism as the sole metaphysical theory of consciousness.[95]

On the contrary, Beni, who begins with a structuralist theory of consciousness and selfhood, faces a challenge in defending any form of panpsychism within this structuralist framework. However, Beni can embrace Mindt's proposed solutions regarding different structures and dynamics. I propose that Beni should align with the trajectory of Mindt's solutions and incorporate these novel understandings of structure and information into his framework. Beni's structuralist theory can readily integrate the differentiation between extrinsic and intrinsic structure and dynamics, along with the associated concepts of intrinsic structure and intrinsic perspective. Given that Beni is already utilizing the free-energy framework and Integrated Information Theory, which includes the intrinsic perspective, his theory can easily

[95] Physicalism can also be viewed as positing intrinsic/categorical physical properties. This kind of physicalism is not an option for a structuralist theory. There are varieties of physicalism without intrinsicality — e.g. *Physical Structuralism* (Ney, 2015).

accommodate the types of information postulated by Mindt. Beni, who does not distinguish between various information concepts and has overlooked the significance of intrinsic perspective in IIT, could enhance his theory by providing a more nuanced portrayal of different forms of information. Clarifying what sets selves apart from other structures would be crucial for Beni, and the introduced notions can help address this issue and potentially develop solutions.

This entails adopting a broader understanding of structure and dynamics, extending beyond the narrow conception of the physical and the natural world (extrinsic conception of structure and dynamics).[96] Making these distinctions aids in maintaining the differentiation between the physical and the mental while recognizing both as structures. With these new notions in play, a clear demarcation between what is physical, what is alive, and what is conscious becomes possible.

As previously discussed, the metaphysical stance of physicalism relies on physical theory, describing the natural world through structural and dynamical processes/properties, termed extrinsic features by Mindt. Postulating intrinsic information and structure already transcends the confines of the "physical" and fully describes systems using traditional concepts from physical sciences alone. The position that appears more fitting for the structuralist aiming to address the hard problem of consciousness is *neutral monism*[97], as hinted earlier. Neutral monism, coupled with an information-theoretic structuralist theory, presents a more promising option for overcoming the hard problem.

[96] On the metaphysical implication of intrinsic structure see Mindt, 2021: 16–17. I side with Mindt that, at this point, physicalism is to be abandoned and neutral monism embraced.

[97] Mindt (Mindt, 2019, 2021) argues that if the intrinsicality/categoricality is denied, then structuralism and neutrality follow.

6.5 NEUTRAL MONIST STRUCTURALISM

Neutral monistic ontology posits that entities we classify as physical and mental share a fundamental essence, suggesting that ultimate reality is singular in nature. Even if we acknowledge distinctions between the mental and the physical, neutral monism remains compatible, allowing for the possibility that diverse entities may derive from a common, ultimate neutral source. While neutral monism appears to provide an elegant resolution to key metaphysical challenges, such as the hard problem of consciousness, it is not without its own complexities. Fundamental questions arise for any adherent of neutral monism, such as the nature of these neutral entities and the relationship between the ultimate neutral and the derived entities of matter and mind.

Traditionally, Mach (Mach, 1886), James (James, 1912) and Russell (Russell, 1927) are considered to be the main proponents of neutral monism. Russell, for example, speaks of the *common ancestor* of mind and matter:

> The stuff of which the world of our experience is composed is, in my belief, neither mind nor matter, but something more primitive than either. Both mind and matter seem to be composite, and the stuff of which they are compounded lies in a sense between the two, in a sense above them both, like a common ancestor. (Russell, 1921: 2)

There have been some examples of combining informational and structural ontologies (Ladyman et al., 2007; Floridi, 2011), but rarely have they been put together with neutral monism (Sayre, 1976; Mindt, 2019, 2021). Sayre advocated a different type of neutral monism from Russell's. Mindt has stated that Sayre, in his version of information-theoretic neutral monism, attempts to accomplish a feat that Russell could not. Mindt (Mindt, 2019: 115) has followed Sayre in criticizing Russell's own neutral monism in that its characterization of the neutral element (neutral "sensibilia") was "useless for any practical purposes" and would have to be rendered back into physical and mental terms. Sayre has put forward an ontological claim that the ultimate nature of reality consists of informational states. Like Sayre, Mindt

argues that the fundamental neutral element is information, applicable both to the physical and the mental domain. Mindt is, therefore, committed to the "Neither View" of neutrality, meaning that he views the fundamental entities as neither physical nor mental, but a third category of entities.[98]

Ladyman and colleagues articulated the idea of information-theoretic structural realism (ITSR), which Beni incorporates into his theory, as previously mentioned. In ITSR, the argument is that our scientific inquiries operate across various scales, unveiling the structural features inherent in reality. These identified patterns are considered ontologically real and exist independently in the world. However, a potential limitation arises with ITSR, labelled as "weak unification" metaphysically due to empirical evidence (Mindt, 2019: 122; Ladyman et al., 2007: 290). To address this limitation, Mindt proposes pushing the information-theoretic ontology further by coupling it with an *information-theoretic neutral monism*. This involves positing that neutral entities constitute the fundamental fabric of reality. Mindt argues in favour of this perspective abductively, asserting that the fundamental nature of reality is information-theoretic, understood as neutral and structural (Mindt, 2019: 123). The resulting position is termed *information-theoretic neutral-structuralism* (ITNS).

The dilemma arises: why posit neutral information as fundamental if the ultimate nature of the universe remains elusive? Mindt contends that adopting ITNS has "strong utilitarian reasons" compared to inertly adhering to physicalism. This theory is favoured since it does not only accommodate consciousness within the natural framework but also hold the promise of bridging the gap between the physical and the mental. Additionally, it aligns closely with the insights derived from the structure of our scientific theories (Mindt, 2019: 124).

Given Beni's commitment to grappling with the complexities of consciousness and selfhood and constructing a naturalistic theory of consciousness using structuralist tools, I propose that he should heed Mindt's suggestions. Beni could argue that the

[98] There are several proposals on how to understand neutrality in neutral monism, and *The Neither View* is one option. See Stubenberg, 2018.

fundamental essence of reality is informational, structural, and neutral—embracing a version of neutral-structuralism. I endeavoured in this Chapter to enhance Beni's information-theoretic structuralist theory of consciousness and self by drawing insights from a similar ontological framework articulated by Mindt. Both Beni and Mindt employ an information-theoretic approach to explaining consciousness, utilizing the conceptual tools of Integrated Information Theory.

Expressing doubt about the compatibility of panpsychism with structuralism, I argued that Beni's theory could be enriched by incorporating solutions proposed to amend Integrated Information Theory. Specifically, recognizing different types of structures could enhance the structuralist theory, surpassing the narrow confines of a purely physicalist perspective. Additionally, exploring ideas related to extrinsic and intrinsic structure and dynamics could offer valuable advancements in addressing fundamental issues in neutral monism—clarifying the nature of ultimate neutral entities and their relationship to entities recognized as physical and mental.

The proposition put forth suggested that neutral monism aligns well with the core tenets of structuralism. Combining the structuralist theory of consciousness and the self together with a neutral monist ontology could yield a more viable naturalistic account of consciousness, presenting a perspective with greater plausibility and alleviating the challenges posed by the hard problem. This proposed neutral-structuralist theory of consciousness and selfhood, however, remains tentative and necessitates thorough development in future research.[99]

[99] I explored these solutions in Nešić 2022a. In a response paper, Beni (Beni, 2022) defends his naturalistic tendencies of the SRS which does not commit to intrinsic phenomenal aspects. He argues that since selves cannot be distinguished from other kinds of structures in the universe this pushes him to embrace a limited form of panpsychism, "panpsychism is mainly about the indiscernibility of the selves from one another and their environment" (Beni, 2022: 6). Since there are no commitments to the intrinsic vs. the extrinsic dichotomy, as Beni emphasises, SRS is in line with my neutral monist tendencies in the target paper.

7. ENACTED SELVES

"I stim, therefore I am."

Melanie Yergeau
(YouTube, January 26, 2012)

7.1 THE DANCER AND THE DANCE

In this final chapter, I will consider enactive and computational theories of the self and a possible integration of the two approaches. This brings us to the end of the modern selfhood saga. These theories are probably the best candidates for the naturalistic theory of the self, given how they combine neuroscience and phenomenology. Perhaps, if not at the moment, then in the future, they will provide us with the best answer to the philosophical question of what consciousness and the self are.

Two embodied approaches to cognition are *enactivism* and *extended cognition* (Di Paolo, 2009; de Haan, 2020b: 46). The thesis on extended cognition originates from the seminal work of Clark and Chalmers (Clark & Chalmers, 1998), who introduced this term. They claim that consciousness can be extended outside the skull so that parts of the world can be carriers of psychic processes at a certain moment. The environment can become part of cognitive processes, such as when we use pen and paper to perform complicated mathematical calculations. Does this mean that consciousness is extended into the environment? A famous example, cited so many times in the literature, is that of Otto. Otto has Alzheimer's disease and has to write down what he wants to remember in his notebook. Clark and Chalmers argue that Otto's notebook contains instantiations of his beliefs, given the role that information in the notebook plays. So, at least when it comes to beliefs, they can be partly constituted by the

environment outside our body. Sannake de Haan points us to an important distinction between the approach to cognition as extended and enactivism. The thesis of extended cognition remains on the ground of internalism, and it is claimed that there is a clear boundary between the inner consciousness/mind and the external world, although this boundary can be moved outwards. In enactivism, the border is completely erased because enactivism calls into question the very dichotomy of internal/external when it comes to consciousness and cognition (Di Paolo, 2009; de Haan, 2020b).

Francisco Varela, Evan Thompson and Eleanor Rosh laid the foundations of enactivism in their book *The Embodied Mind* (Varela et al., 1991). This position was put forward as a reaction to the cognitivist understanding of consciousness/mind. In cognitivism, the human mind is like a computer system that operates on mental representations of the external world according to certain rules. There is a clear separation of the mind as internal and the world/environment as external, and the mental processes are in the brain.[100] Enactivism calls into question such an image of the human brain, consciousness and cognition and advocates the understanding of consciousness and cognition as non-representational, denies that they are closed in the brain, but rather claims that cognition is embodied action.[101] Enactivists deny that there is a separation of internal and external, which means that the organism and the world/environment are dynamically coupled (dynamically coupled; de Haan, 2020b: 53).

Enactivism is a very broad research program, not just one theory, and we can talk about three different approaches within this movement. Classical enactivism (the earliest form of enactivism) connects biological explanations of the origin of life with explanations of sensorimotor processes, sensorimotor enactivism focuses on perception, while radical enactivism (Hutto & Myin,

[100] Ideas about embodied and immersed consciousness have existed since the beginning of the twentieth century, for an overview see de Haan, 2020b: 52.

[101] Theories of predictive coding and predictive processing, in their traditional form, remain representational (Hohwy, 2016), but there are also those who advocate that these theories can be made compatible with enactivism (Bruineberg & Rietveld, 2014; Bruineberg et al., 2018; cf. Di Paolo, Thompson & Beer, 2022).

2013) rejects any version of representationalism (the enactive approach, sensorimotor enactivism, and radical enactivism; Heras-Escribano, 2021: 343).

The enactivist approach in cognitive science is a version of the embodied cognition paradigm (Thompson, 2007). According to enactivism, a living organism is not an isolated individual but part of a dynamic system together with its environment, and this system must be taken into account when studying cognition. Cognition is embodied and shared, meaning that the organism learns about the world through activity through interaction with the environment. Cognition is exactly that activity (interaction) of an organism that tries to understand its environment — *sense-making,* in enactivist terminology. What the organism encounters in the environment is either bad for it, scares it (when we see a poisonous spider) or is attractive (when we feel the warm touch of a close person). Through interaction, the organism bodily evaluates the circumstances in its environment and evaluates the affective attractions (or repulsions) of the situations in which it finds itself (de Haan, 2020a, 2020b).

Enactivists wanted to reconcile cognitive science with the phenomenology of everyday experience. To solve the problem of consciousness (Chalmers, 1996) and reduce the gap between nature and consciousness, enactivists build their research program on ideas from biology, phenomenology and neuroscience (Thompson, 2007). Even in the first book, they refer to the phenomenological theories of Husserl, Heidegger and Merleau-Ponty, but they criticize Western phenomenology as a failed philosophical project. In later works and books, Evan Thompson will try to correct this injustice to phenomenology (especially to Husserl, whose views he admits to having misinterpreted) and defend the position that phenomenology can be of great importance to cognitive science (Thompson, 2007).[102] Varela, with his "neurophenomenology", was the first to create a framework in which neuroscience and phenomenology are integrated (Varela, 1996).

Enactivists advocate the thesis about the continuity of life and consciousness/mind (*life-mind continuity thesis*) (Di Paolo,

[102] Similarly in Gallagher & Zahavi, 2008.

2009; Froese & Di Paolo, 2009; Thompson, 2007). Simply put, where there is life, there is consciousness. For enactivists, this means that where there is life, there is also meaningful activity in the adaptive interaction of the organism with the environment. Life and biological organisms are characterized by the ability to self-maintain and self-organize (autopoiesis). The term autopoiesis was introduced by Varela and Maturana (Varela & Maturana, 1972/1980) to explain how the autonomy of living things is organized. Given that there is a gradation in system survival (it doesn't have to be all-or-nothing), Di Paolo (Di Paolo, 2005) adds the concept of system adaptability and defines an organism as a "network of processes" that is separated from the environment but is in constant interaction with it. Living beings as *autonomous systems* have the properties of *operational closure* and *precariousness* (Maturana & Varela, 1987; Kiverstein, 2018).

An autopoietic system is one that is continuously self-producing, a system that autonomously creates its own identity (Froese & Di Paolo, 2009). Living organisms constantly strive for self-preservation; "life says yes to itself" (Jonas, 1992: 36; cited in Froese & Di Paolo, 2009). Maintaining identity is central to a living organism, and values for that organism emerge in the context of interaction with the environment — the organism's interaction with the environment is normative (Froese & Di Paolo, 2009: 6; Di Paolo, 2009: 12). The autopoietic system, therefore, is purposeful in two ways, by creating and maintaining a dynamic identity through change and by evaluating the environment in order to maintain it (*sense-making*; Thompson, 2007: 146–147). The environment has meaning and meaning for the organism (*Umwelt*; von Uexküll, 1909). It is not just a set of physical and chemical properties but a meaningful ecological niche, the living environment of that organism. According to enactivism, the organism gives meaning to the environment and projects meaning into the environment.

Life and consciousness/mind share common organizational properties. The problem is how to explain that the same principles can apply both to the simplest organisms and to the highest forms of human cognition ("cognitive gap", De Jaegher & Froese, 2009). Enactivists argue that the problem of the cognitive gap

can only be solved when the constitutive role of intersubjectivity in these processes is considered.

According to enactivists, the mind/consciousness is an embodied dynamic system. In accordance with this dynamic understanding of life and consciousness, they usually represent emergence, and the type of causality most often mentioned in such systems is circular causality (Thompson & Varela, 2001; Fuchs, 2018). As an autopoietic system, a living organism is an autonomous system that creates and maintains its own identity. In the autonomy of this system, we see the emergence of dynamic co-emergence; the whole emerges from the parts, and the parts emerge from the whole; the parts and the whole co-emerge and determine each other (Thompson, 2007: 65). Enactivists refer to Merleau-Ponty, who argued that the relationship between an organism and its environment (fr. *milieu*) is not described by linear causality, but rather circular causality (Merleau-Ponty, 1963). Thompson talks about emergent processes, not properties. What emerges is the self or the individual as a process, together with its niche (environment). One of the targets of de Haan's criticism is the use of downward causation in enactivism, which seems to invoke traditional dualism (mental and physical) and the existence of levels. Thompson (Thompson, 2007) admits that downward causation is an inappropriate metaphor for describing the influence of the whole on its parts and, therefore, advocates for relational holism in which there are no upper and lower layers.

Enactivists often understand causality in a living organism as circular, with vertical and horizontal dimensions: vertically — up and down within the organism (bottom-up and top-down causality), and horizontally — as interactions with the environment (Fuchs, 2018: subsection 3.3). De Haan criticises this understanding of causality because she believes that in this way, processes are still separated and they should not be, and that the existence of separated layers (levels) is implied (de Haan, 2020b: subsection 4.7.2). There is no causality between the experiential and physiological because they are not separate processes and properties (as in dualism) but are part of the person-world system.

According to enactivism, perception and action (sensory and motor processes) are inseparable in cognition, so perception

is also an activity, not a passive creation of internal representations about the world (Varela et al., 1991; Noë, 2004). However, enactivism applies to all mental processes, not just perception and action. There are also enactivist approaches to intersubjectivity and social cognition (De Jaegher & Di Paolo, 2007; Fuchs & De Jaegher, 2009), as well as language (Di Paolo et al., 2018), and it is also applied in the research of human cognition in the deep past and archaeology (Malafouris, 2013).

General applications of embodied and enactivist ideas in psychiatry and psychopathology already exist (Hutto, 2010; Fuchs, 2018; Myin et al., 2015; Maiese, 2016). Enactivism has been invoked to aid the understanding of autism (De Jaegher, 2013; Klin, Jones, Schultz, & Volkmar, 2003)[103] and schizophrenia (Kyselo, 2016; Krueger, 2020). Also, it should be noted that the extended approach to cognition has also been applied to psychiatric disorders (Hoffman, 2016; Krueger, 2020; Roberts, Krueger & Glackin, 2019; de Haan, 2020b).[104]

I mentioned one enactivist theory of the self in Chapter 6, Gallagher's self-pattern theory. This is an approach that avoids understanding the self as both an illusion or a substance and falls somewhere between the traditions of Descartes on one side and Metzinger on the other. It is neither reductionist nor deflationist and more akin to the pluralistic views of the self of James (James, 1890) and Neisser (Neisser, 1988). Gallagher puts forward his theory in an attempt to overcome reductionism and take into account all the diverse aspects of the self, be they biological or phenomenological. This kind of pluralist approach promises to answer the worry that there are too many radically different conceptions of selfhood, both in philosophy and science and no unificatory framework in sight. According to the self-pattern theory, what we call the self "is a pattern of dynamically integrated processes or factors" (Gallagher, 2013; Gallagher et al., 2023; Thompson, 2020).[105] The factors that weave themselves

[103] I defended an enactive and ecological account of autism in Nešić, 2023a.

[104] For more on the enactivist applications in psychiatry, see Nešić, 2022b.

[105] These are "real patterns" from the position of positive realist approaches, as discussed in Chapter 6. See Dennett 1991 and Ladyman & Ross, 2007. Like the structuralist theory of the self, this is neither a substantivalist nor no-self view

into the self-pattern are bodily, agentive, interpersonal, narrative and ecological. The self is not just an aggregate of these factors but a *dynamical gestalt* (Gallagher, 2021), with a set of relations between those factors; the self is a recursively self-organizing pattern (Varela, 1997). So, the self is a system of processes with only several jointly sufficient conditions. Thus, even persons with severe Korsakoff syndrome (dysnarrativa) can still be considered to have a self, where although some major aspects or factors may be missing, some others are preserved. What becomes part of the self-pattern are the following processes: bodily processes, pre-reflective experiential processes, affective, behavioural/action-related, social/intersubjective, cognitive and psychological processes, reflective, narrative, ecological and normative processes (Gallagher, 2021: 129, 2023).

The self-pattern has enactive characteristics; it is dynamically changed by factors of the whole brain-body-environment system (this is already evident in the set of factors that make up the self-pattern). Gallagher and colleagues see this pattern theory of the self as complementary to the Buddhist psychology of the self.[106]

Gallagher offers the same solution for the integration problem in psychiatry. De Haan has pointed out that there is a problem of integration in psychiatry — how to connect very different factors that contribute to the emergence and

and could be fitted into the bundle self category. Gallagher himself points out that it is neither a substance, nor the no-self view. If a self is a structure, it is real; it is something, a pattern, a process, and not a permanent substance. Still, structures have no thick properties, so maybe it could not even be called a bundle; this is more deflationist. Generally speaking, if enactive self theories are to be considered as bundlist, it will depend on the ontology behind enactivism. Since there is no explicit metaphysics in enactivism, at least not in terms of the Western analytic philosophical tradition, it would seem imprecise to classify such theories of the self as bundlist. Varela (Varela, 1996) intended neurophenomenology to be a *methodological* remedy for the hard problem of consciousness (Chalmers, 1996). Still, those like Gallagher try to situate their enactive views in modern metaphysics. See Pace 2021a for some discussion on these matters. He complements phenomenology and enactivism with a kind of neutral monism (Pace, 2021b)

[106] Although Buddhism is usually viewed as a position in which the existence of the self is denied (the non-self view), the situation seems to be a bit more complicated, and this ontology needs unpacking. For this, see papers in Siderits, Thompson and Zahavi, 2011 and Thompson, 2020.

development of psychiatric disorders? Usually, disorders are discussed from a neurophysiological perspective or from a social or phenomenological (subjective) perspective. The question arises whether one of these perspectives alone is sufficient for a complete explanation of psychiatric disorders.

De Haan finds weaknesses in contemporary models in psychiatry. She characterized and divided these models into 1) reductionist, 2) dualistic, and 3) integrative (de Haan, 2020b: section 2). One-sided ones, such as biological models and neuroreductionism (which is the most frequently used model in psychiatry), view these disorders as brain diseases whose causes can be found in genetics and neurophysiological processes. However, social, phenomenological, psychoanalytic, and existential models are also one-sided. Such models are integrative only in the sense that they reduce all factors to one dimension. There are also dualistic models, which combine scientific facts with patient experiences (two-sided models). Integrative models, as the biopsychosocial model (Engel, 1977), try to connect different perspectives. Similarly, Sass et al. (Sass et al., 2018) give a biophenosocial model that distinguishes levels, but cannot answer how these aspects are connected. De Haan precisely criticizes these models because they do not show us how different aspects and factors are related. She believes that there are four dimensions (aspects) that every integrative model must include: experiential, physiological, sociocultural, and existential dimensions (de Haan, 2020a: 11).

De Haan proposes a model that is within the framework of enactivism, but unlike previous models, it is not based on hierarchical level-based explanations. Her model for psychiatry is based on *organizational causality* in which causal relations are neither top-down nor bottom-up, but are organized holistically and hierarchically (Gallagher, 2022: 2). In this kind of causation, we can distinguish between global and local processes, which allow us to determine the causes of psychiatric disorders more precisely. In de Haan's model, emergence is a fusion process, which means that certain processes and factors disappear and merge. Gallagher rightly criticizes the introduction of this understanding of causation and integration because, in the psychiatric context, we want to connect different factors meaningfully but in such a way

as to keep the distinctions between those factors, like the four dimensions that de Haan also distinguishes (Gallagher, 2022: 3). To avoid these problems, Gallagher proposes the idea of *dynamic causality* and of the *dynamic gestalt*, where processes, which can still be distinguished, are dynamically connected (this gestalt is heterarchical and without levels). According to Gallagher, the dynamic gestalt is a better solution because, unlike the usual gestalt, where we have a whole that is more than the sum of its parts, this type of gestalt does not consist of relations between the whole and the parts but of dynamic causal relations that form a pattern. This would also be important for diagnosis in psychiatry. The symptoms of psychiatric disorders do not occur in isolation but can be seen as wholes of which various symptoms are only aspects.

Myriam Kyselo (Kyselo, 2014) has acknowledged that Gallagher's pattern theory of the self makes progress in addressing as many different aspects or factors of the self as possible but has warned that it lacks the integration it strives for. Its openness (why any of the factors figure in and how do they relate) makes it prone to such criticism because it gives no definite answer to how we get to the whole (the *individual self*)[107]. Moreover, Kyselo rightly emphasises that this is not just an armchair problem but a methodological and practical one for all those involved in the research of the self (cognitive scientists), as well as for those diagnosing psychopathologies of the self and finding proper ways of treating them, like psychiatrists (Kyselo, 2014: 2).

In order to solve adequately the problem she finds in Gallagher's account, Kyselo proposes that the enactive notion of autonomy will be helpful for understanding the individuation and unity of the self. A related problem that she dubs the *body-social problem*[108] is something that needs to be tackled. The self and the body are still, despite the interactional and embodied turn in philosophy and science,[109] considered way too

[107] This was a desideratum when it comes to an understanding the self, which Olson pointed out long ago (Olson, 1998).

[108] Alluding to the *body-mind* problem.

[109] We have been witnessing an *interactive turn* in philosophy and cognitive science for some time. This movement seems to be a part of a larger *intersubjective* turn

individualistically without the thorough incorporation of the social. The problem relates to how both the body and the social factors shape the individuation of the self. This divide between selves has to be bridged. Her enactive theory of the self builds on the pluralistic approach while avoiding the pitfalls of essentialism (the self being solely bodily or social). However, even on the embodied approach to cognition, there seems to be a core, bodily self, independent of the social (Zahavi, 2014) and the self, although embedded, is individualized on its own and Kyselo calls this claim *social as contextual*. On the other side, there is a strong *social as constitutive* claim to which only some would subscribe (de Haan, 2010).[110]

Using the idea of autonomy, Kyselo defends an enactive account of the self in which the principle of individuation is based on Jonas's concept of "needful freedom" (Jonas, 1966/2001), where an organism individuates itself *through and from a world*. It needs the world but also has to be emancipated from it. This principle of individuation can be extended to the social, and the self is defined organizationally as a whole through social and environmental interaction. The self is not given but open and ever-changing. Human self identity takes place in "social needful freedom". The individual is a *self-other generated autonomous network*. Finally, the autonomy of the individual is defined by Kyselo as such:

in the embodied, embedded, and enactive cognitive science of the mind (Varela, Thompson & Rosch, 1991; Thompson, 2007; De Jaegher, 2018). A scientific approach, though it still seems to be an exception rather than a rule in neuroscience, is *second-person neuroscience* (Bolis & Schilbach, 2020; Schilbach et al., 2013), which studies social cognition through understanding social interaction. Here, neural mechanisms of social cognition/interaction are considered in a decidedly second-personal/interactive rather than an observational (third-personal) context.

[110] At the time this article was written, perhaps. Since then, more and more philosophers have claimed that even the minimal self is relational and socially constituted. See Ciaunica 2016, Ciaunica & Crucianelli 2019, and the papers in the special issue, Ciaunica 2020. Still, Kyselo's claim is strictly about enactivism. She criticizes Zahavi's view on the relation betwixt the minimal self and the social self in Kyselo (Kyselo, 2016). According to Zahavi (who is inspired by Husserl), there is a special kind of personal self-consciousness that comes between minimal, pre-reflective self-awareness and full reflective self-awareness. A common world is created between communicating subjects and mutual understanding. Each of those subjects must reciprocally agree to respond to the other subject.

> "Individual autonomy is a self-other generated network of precariously organized interpersonal processes whose systemic identity emerges as a result of a continuous engagement in social interactions and relations that can be qualified as moving in two opposed directions, toward emancipation from others (distinction) and toward openness to them (participation)." (Kyselo, 2014: 10)

She argues that the self is not the body; the body is a mediator for the self, a "sensor" for monitoring social interactions. The human self is a "socially co-enacted identity", and in her view even bodily self-awareness emerges through relations with others. Bodily consciousness in itself is not enough even for the minimal sense of human self, Kyselo is adamant. She also provides several empirical examples (e.g. locked-in syndrome, Moebius syndrome, social pain) to back up her claims on how the social is constitutively important for the self.

In a different paper, Kyselo, in a detailed manner, emphasizes how Varela in his work gave the conceptual foundation for the enactive theory of the self.[111] The enactive view is that "the self (in the sense of diachronic and synchronic identity) continuously emerges from various types of intertwined organismic activities" (Kyselo, 2023: 1; Varela, 1991). There is a bodily sense of the unity of self, but this unity is not substantival but *processual* in nature. This idea of the self has been perpetuated throughout enactivism: "the self *is* the sum of the autonomous network's invariant self-organising activities" (Kyselo, 2023: 2; Thompson, 2007: 61).

Thompson, Gallagher, Kyselo, all fall back on Varela's foundational insights and follow this tradition of enactive selfhood.[112] Varela understands the self through emergent properties in the distributed network processes and calls an organism a *meshwork of selfless selves*: "Organisms, those fascinating meshworks of selfless selves, no more nor less than open-ended, multi-level circular

[111] In this article, Kyselo argues that Varela in his early paper "Not one, not two" makes a "paradigm shift" towards the social self view that she also defends.

[112] The same no-ego-self view based on both cognitive science and Buddhist phenomenology (mindfulness meditation) was expounded in Varela, Thompson and Rosch (Varela, Thompson & Rosch, 1991: Chapter 4).

existences, always driven by the lack of significance they engender by asserting their presence" (Varela, 1991: 104).

Thompson has expressed this idea of the enactive self in a more poetic manner:

> "The way I like to put this idea is that a self is an ongoing process that enacts an "I" and in which the "I" is no different from the process itself, rather like the way dancing is a process that enacts a dance and in which the dance is no different from the dancing" (Thompson, 2020, 2015).

7.2 UNCERTAIN SELVES

The Free Energy Principle (FEP) holds the potential to serve as a unifying theory in the realms of biological and cognitive sciences. In maintaining its organization as an adaptive living system, an organism strategically minimizes its information-theoretic free energy through interactions with its environment (Friston & Stephan, 2007).[113] This minimization is achieved either by predicting sensory input, or by modifying the environment to align with the predictions, encompassing both perceptual and active inference as means to bridge models and the world. This process allows an organism to finely tune itself to its ecological niche.

Bruineberg et al. underscore that FEP serves as a comprehensive framework for self-organizing living systems, with predictive coding/processing focusing on the neural functioning of the brain. While FEP and Bayesian predictive coding often coexist, the authors argue against conflating them, emphasizing their individual significance (Bruineberg, Kiverstein & Rietveld, 2018: 2419).

Predictive processing, or prediction-error minimization, constitutes a theory elucidating the functions of the brain and its cognitive processes (Clark, 2013; Hohwy, 2013; Friston, 2010). Recently, it has been employed as a theoretical foundation for exploring mental illness within computational psychiatry. The predictive brain strives to minimize prediction-errors arising from the

[113] For explanations of complicated terminology of predictive processing/coding and the Free Energy Principle, see Bruineberg, Kiverstein and Rietveld 2018.

mismatches between top-down predictions and bottom-up sensory information. The brain essentially embodies a hierarchical probabilistic model of the environment referred to as the generative model. Agents allocate varying levels of precision to prior beliefs or current sensory evidence (prediction errors) based on their perceived reliability, or "precision." In this context, perceptual and active inference are not distinct strategies for minimizing prediction-error but are regarded as integral components of a unified process, preparing the organism to act in ways that enhance its well-being (Bruineberg, Kiverstein & Rietveld, 2018: 2430).

A distinction can be made between deflationary and non-deflationary accounts of the self within predictive processing and active inference frameworks (Perrykkad & Hohwy, 2020). A deflationary one in which the self "falls out of the process of active prediction error minimisation" and is not explicitly represented within the system is Friston's view (Friston, 2018). Close to Friston's is Kiverstein's view (Kiverstein, 2020). Less deflationary are those accounts that see the self as an inferred hidden cause of sensory experience (Apps & Tsakiris, 2014; Letheby & Gerrans, 2017). Through the process of prediction-error minimization, the organism gives "constant evidence for its own existence". In this interpretation, the self is an inferred network of causes (Hohwy & Michael, 2017), so one can take several stances towards the reality of the self as the referent of the model: sceptical, agnostic, realist. Realists think the self-model is the self (e.g. Hohwy & Michael, 2017; Limanowski & Blankenburg, 2013). Perrykkad and Hohwy (Perrykkad & Hohwy, 2020) and base their perspectives on the predictive processing accounts of the self in cognitive neuroscience, particularly drawing from Apps and Tsakiris and Limanowski and Blankenburg (Limanowski & Blankenburg, 2013).[114] As expected, Metzinger (Metzinger, 2003) was the forerunner for modern views on the self-model.

These are the main PP approaches to the minimal self, which is usually perceived as the bodily self. Apps and Tsakiris attempt to explain bodily self-awareness through PP, while

[114] So do Constant, Bervoets et al. 2018 when they explore the relational self in autism through PP accounts. I tackle these issues surrounding autism in Nešić, 2023a.

Limanowski and Blankenburg connect the predictive processing framework with minimal selfhood, and Seth (Seth, 2013) has argued that understanding the internal state representations (interoception and proprioception) for self-representation. According to Seth, an active inference on internal states gives us the phenomenology of body ownership *(body as an object*, Perrykkad & Hohwy, 2020). He says: "Experience of body ownership, a key aspect of selfhood, is modulated by predictive multisensory integration of precision-weighted interoceptive and exteroceptive signals" (Seth, 2013: 571). This experience of being somebody, of *being you*, emerges from the brain's predictions of the internal bodily states; it is a "controlled hallucination" (Seth, 2021).

Building on traditional predictive coding/predictive processing theories, embodied and enactive interpretations of PP and active inference have surfaced. Integrating computational and embodied approaches to cognition, consciousness, and the self has not been an easy task and has not always been met with enthusiasm. One example of a good integration that I find very useful for thinking about the self and consciousness, which I will consider in this chapter is the *Skilled Intentionality Framework* (SIF). SIF incorporates an *ecological-enactive interpretation* of FEP (Bruineberg & Rietveld, 2014; Bruineberg, Kiverstein & Rietveld, 2018). The ecological and enactive framework (Rietveld, Denys & van Westen, 2018) brings together complementary findings across various scientific disciplines, encompassing neurodynamic, ecological, affective, and phenomenological analyses of cognition—all within the framework of the self-organizing system "brain–body–landscape of affordances," considering both individual and environmental factors. SIF establishes connections among diverse disciplines, including ecological psychology (involving the landscape of affordances; Gibson, 1979; Heft, 2001; Chemero, 2009), phenomenology (addressing selective openness to and relevance of affordances, as well as optimal grip), emotion psychology (exploring states of action-readiness akin to Frijda, 2007), and embodied neurodynamics (examining self-organizing affordance-related states of action-readiness).

Within SIF, an integration of embodied, enactive, and ecological perspectives, cognition is conceptualized as adept

involvement with various affordances (opportunities for action) within the sociomaterial environment of an individual's ecological niche. This engagement is characterized by a tendency toward achieving an optimal grip. An integral aspect of the SIF involves an ecological-enactive interpretation of the free energy principle and predictive processing, as Bruineberg and Rietveld outlined.

At the heart of the phenomenological perspective is the pivotal concept of the optimal grip, rooted in Merleau-Ponty's philosophy of life (Merleau-Ponty, 1968/2003), in which he posits that all living beings inherently exist in a state of disequilibrium within the individual-environment system. This inherent lack serves as a driving force for what he terms "compensatory activity" (Merleau-Ponty, 1968/2003: 149; Rietveld, 2008: Ch. 7), and the experienced result is an "affective tension."

To grasp the notion of the optimal grip, consider Merleau-Ponty's well-known example of approaching a painting in an art gallery to find the ideal viewing distance (Merleau-Ponty, 1945/2002). This illustrates why living organisms are continually selectively open to the landscape of affordances, responding to relevant opportunities (Bruineberg & Rietveld, 2014). In the dynamic coupling of body and world, organisms consistently strive for an optimal grip. Being responsive to solicitations becomes crucial for individuals to enhance their situation.

In essence, the core tenet of Skilled Intentionality can be succinctly expressed as: "Skilled intentionality entails reducing disequilibrium by simultaneously moving toward an optimal grip on multiple relevant affordances, encapsulated within a field of relevant affordances" (Rietveld, Denys & van Westen, 2018: 45).

Advocates of SIF view the free energy principle as asserting that the brain is an integral part of a larger interconnected system with the environment. This interconnected system continually minimizes misattunement with the environment, striving for an optimal grasp on relevant affordances. SIF posits that this optimization reduces disequilibrium within the dynamic system of "brain–body–landscape of affordances." By minimizing free energy, the internal dynamics of the brain are naturally attuned to the external dynamics of the environment (Bruineberg, Kiverstein & Rietveld, 2018: 2440).

Following this perspective, Kiverstein (Kiverstein, 2018) advances an ecological and enactive understanding of the self in an active inference framework (under the SIF). He offers a naturalistic explanation of minimal selfhood as emergent from self-organising biological processes, specifically in processes of prediction-error minimisation. The answer Kiverstein gives to the question of when the subjective life begins for organic life is that it arises when processes of purposive agency and sensorimotor integration are combined, neither of them being sufficient for subjectivity on their own. He follows Friston in claiming that once an organism reaches enough complexity and act so that it minimises its own free energy, "mineness" emerges as intrinsic to the living system (Kiverstein, 2020; Friston, 2018).

Kiverstein sees the organism as biologically embodying a model of its own existence; it is a model of itself in the econiche, a self-model. This is not enough for a self; what needs to be added to the self-model for an organism to become a self is mineness, which is achieved through purposive agency and sensorimotor integration. Although, in agreement with some other PP and active inference accounts of the self, Kiverstein does not posit that the self is a product of the brain's abductive inferences (minimal self, with mineness). If the self is probabilistically inferred, we can be wrong about ourselves in all possible ways, which has been a worry inherent in many views about the self from the very beginning of this book (and something that seems to be the case in many psychopathologies). Kiverstein wants to maintain the phenomenological stance that views mineness as intrinsic to life and organisms of enough complexity. He wants to preserve the phenomenological claim that experiences are logically immune to error through misidentification.[115] He is eager to argue that his ecological-enactive interpretation of active inference ensures no such consequence.

In Kiverstein's account, a self-model is conceived as the agent's representation of its selective engagement with affordances (Kiverstein, 2018: 7). Through active inference, involving

[115] In the case of the *rubber hand* illusion, Kiverstein points out that what is investigated here is the awareness of the "body as object" and not as a subject, and hence can be prone to error (Kiverstein, 2018: 13).

cycles of perception and action, the entire organism adjusts its dynamic coupling to the environment, maintaining operational closure across various organizational levels (Kiverstein, 2018: 9). These systems are deemed "self-specifying" due to the systematic relationship between sensing and moving realized through the perception–action cycle, encompassing both perceptual states (sensorimotor integration) and purposive agency. However, according to Kiverstein, this alone does not account for mineness, a fundamental aspect of self-awareness. He posits that additional *temporally thick* self-models are necessary for subjectivity, drawing on Friston's concept.[116] Kiverstein's understanding of mineness aligns with a relation theory of the self. In an ecological and enactive interpretation of active inference, he emphasizes that "the organism and its environment are co-specifying and co-determining" (Kiverstein, 2020: 3; Gibson, 1979: 4), implying a co-determination between the self and the other.

[116] Temporally thick self-models involve the depth of prior beliefs regarding the achievable future, specifically, beliefs about the repercussions of committing to particular courses of action (where these beliefs are understood strictly in a Bayesian sense, as subpersonal). While these kinds of beliefs are agential and pertain to the self, they are fundamentally prior beliefs and are susceptible to distortion if maintained with excessively low precision or subpersonal conviction. This distortion is precisely the anomaly caused by overly precise sensory precision, wherein the likelihood component of the generative model is endowed with excessive precision. The precision-related anomaly contributing to Autism Spectrum Disorder (ASD), for example, unavoidably limits the depth or time horizon of any planned interaction with the world (or the body), whether these plans involve the interoceptive or prosocial domains. I have written about the autistic self in Nešić 2023a, 2023c.

REFERENCES

Aaronson, S. (2014). Why I Am Not An Integrated Information Theorist (or, the Unconscious Expander). http://www.scottaaronson.com/blog/?p=1799 (accessed June 14, 2016).

Alter, T. & Nagasawa, Y. (2015). *Consciousness in the Physical World: Perspectives on Russellian Monism*. Oxford: Oxford University Press.

Alter, T. (2016). The Structure and Dynamics Argument against Materialism. *Noûs,* 50, 794–815.

Apps, M. A. & Tsakiris, M. (2014). The free-energy self: a predictive coding account of self-recognition. *Neuroscience and Biobehavioral Reviews*, 41, 85–97.

Barbaras, R. (1991). *De l'être du phénomène: sur l'ontologie de Merleau Ponty*. Grenoble: Diffusion Harmonia Mundi.

Bateson, G. (1972). *Steps to an Ecology of Mind* (1st ed.). San Fransisco: Chandler Pub. Co.

Basile, P. (2010). It Must Be True — but How Can It Be? Some Remarks on Panpsychism and Mental Composition. *Royal Institute of Philosophy Supplement,* 85, 93–112.

Broad, C. D. (1925). *The Mind and Its Place in Nature*. London: Kegan Paul, Trench, Trubner & Co.

Beni, M. D. (2019). *Structuring the self. Series New Directions in Philosophy and Cognitive Science*. Switzerland: Palgrave Macmillan.

Beni, M. D. (2021a). A critical analysis of Markovian monism. *Synthese,* 199, 6407–6427.

Beni, M. D. (2021b). A free energy reconstruction of arguments for panpsychism. *Phenomenology and the Cognitive Sciences,* 22, 399–416,

Beni, M. D. (2022). The Curious Incident of Indistinguishable Selves, A Reply to Nešić. *International Studies in the Philosophy of Science*, 35(3–4), 261–268.

Billon, A. (2017). Basic Self-Awareness. *European Journal of Philosophy,* 25(3), 732–763.

Chalmers, D. (2015). Panpsychism and panprotopsychism. In: T. Alter & Y. Nagasawa (Eds.), *Consciousness in the Physical World: Perspectives on Russellian Monism* (pp. 246–277). Oxford: Oxford University Press.

Bird, A. (2016). Scientific progress. In: P. Humphreys (Ed.), *Oxford handbook in philosophy of science* (pp. 544–565). Oxford: Oxford University Press.

Bolis, D. & Schilbach, L. (2020). "Through others we become ourselves": The dialectics of predictive coding and active inference. *Behavioral and Brain Sciences*, 43(90), 25–26.

Brüntrup, G. (2011). Panpsychism and structural realism. In: M. Blamauer (Ed.), *The mental as fundamental: new perspectives on panpsychism* (pp. 15–35). Frankfurt: Ontos Verlag.

Brüntrup, G. & Jaskolla, L. (eds.) (2016). *Panpsychism: contemporary Perspectives*. New York: Oxford University Press.

Bruineberg, J. & Rietveld, E. (2014). Self-organization, free energy minimization, and optimal grip on a field of affordances. *Frontiers in Human Neuroscience*, 8, 1–14.

Bruineberg, J., Kiverstein, J. & Rietveld, E. (2018). The anticipating brain is not a scientist: the free-energy principle from an ecological-enactive perspective. *Synthese,* 195(6), 2417–2444.

Cartwright, N. (1994). Fundamentalism Vs. The Patchwork of Laws. *Proceedings of the Aristotelian Society,* 94, 279–292.

Chakravartty, A. (2007). *A Metaphysics for Scientific Realism*. Cambridge: Cambridge University Press.

Cerullo, M. A. (2015). The Problem with Phi: A Critique of Integrated Information Theory. *PloS Computational Biology,* 11(9), 1–12.

Chalmers, D. (1995). Facing up to the Problem of Consciousness. *Journal of Consciouness Studies,* 2, 200–219.

Chalmers, D. (1996). *The conscious mind: in search of a fundamental theory*. New York: Oxford University Press.

Chalmers, D. (2003). Consciousness and Its Place in Nature. In: S. P. Stich & T. A. Warfield (Eds.), *The Blackwell Guide to Philosophy of Mind* (pp. 102–142). Blackwell Philosophy Guides. Malden, MA: Blackwell Publishing.

Chalmers, D. (2010). *The Character of Consciousness*. New York: Oxford University.

Chalmers, D. (2012). *Constructing the World*. Oxford: Oxford University Press.

Chalmers, D. (2016). Panpsychism and Panprotopsychism. In: G. Brüntrup & L. Jaskolla (Eds.), *Panpsychism: contemporary Perspectives* (pp. 19–48). New York: Oxford University Press.

Chemero, A. (2009). *Radical embodied cognitive science*. Cambridge, MA: MIT Press.

Chisholm, R. (1969). On the observability of the self. *Philosophy and Phenomenological Research,* 30(1), 7–21.

Ciaunica, A. (2016). Basic forms of pre-reflective self-consciousness: a developmental perspective. In: S. Miguens, G. Preyer & M. C. Bravo (Eds.), *Pre-reflective consciousness: Sartre and contemporary philosophy of mind* (pp. 422–439). London: Routledge.

Ciaunica, A. & Crucianelli, L. (2019). Minimal self-consciousness from within — a developmental perspective. *Journal of Consciousness Studies,* 26(3–4), 207–226.

Ciaunica, A. (2020). Introduction: The Relational Self: Basic Forms of Self-Awareness. *Topoi,* 39(3), 501–507.

Clark, A. & D. Chalmers (1998). The extended mind. *Analysis,* 58(1), 7–19.

Ciaunica A. & Fotopoulou, K. (2017). The touched self. Psychological and philosophical perspectives on proximal intersubjectivity and the self. In: C. Durt, T. Fuchs & E. Tewes (Eds.), *Embodiment, enaction and culture: investigating the constitution of the shared world* (pp. 173–192). Cambridge: MIT Press.

Clark, A. (2013). Whatever next? Predictive brains, situated agents, and the future of cognitive science. *Behavioral and Brain Sciences*, 36(3), 181–204.

Coleman, S. (2013). The real combination problem: panpsychism, micro-subjects, and emergence. *Erkenntnis,* 79, 19–44.

Constant, A., Bervoets, J., Hens, K. & Van de Cruys, S. (2018). Precise Worlds for Certain Minds: An Ecological Perspective on the Relational Self in Autism. *Topoi,* 39, 611–622.

Dainton, B. (2008). *The Phenomenal Self*. Oxford: Oxford University Press.

Dainton, B. (2011). Review of *Consciousness and its Place in Nature*, *Philosophy and Phenomenological Research,* 83, 238–261.

De Haan, S. (2010). Comment: the minimal self is a social self. In: T. Fuchs, H. C. Sattel & P. Henningsen (Eds.), *The Embodied Self* (pp. 12–17). Stuttgart: Schattauer.

De Haan, S. (2020a). An Enactive Approach to Psychiatry. *Philosophy, Psychiatry & Psychology*, 27(1), 3–25.

De Haan, S. (2020b). *Enactive psychiatry*. Cambridge: Cambridge University Press.

Dellsén, F. (2018). Scientific progress: Four accounts. *Philosophy Compass*, 13(11), 1–10.

Dennett, D. (1989). *The Intentional Stance*. Cambridge, MA.: MIT Press.

De Jaegher, H. & Froese, T. (2009). On the role of social interaction in individual agency. *Adaptive Behavior,* 17(5), 444–460.

De Jaegher, H. & Di Paolo, E. (2007). Participatory sense-making: An enactive approach to social cognition. *Phenomenology and the Cognitive Sciences*, 6(4), 485–507.

De Jaegher, H. (2013). Embodiment and sense-making in autism. *Frontiers in Integrative Neuroscience*, 7, 1–19.

De Jaegher, H. (2018). The intersubjective turn. In: A. Newen, S. Gallagher & L. de Bruin (Eds.), *The Oxford Handbook of 4E Cognition: Embodied, Embedded, Enactive and Extended* (pp. 453–468). Oxford: University Press.

Di Paolo, E. (2009). Extended life. *Topoi*, 28(1), 9–21.

Di Paolo, E., Thompson, E. & Beer, R. (2022). Laying down a forking path: Tensions between enaction and the free energy principle. *Philosophy and the Mind Sciences*, 3, 1–39.

Dennett, D. (1991). Real patterns. *Journal of Philosophy,* 88, 27–51.

Duncan, M. (2015). We Are Acquainted With Ourselves. *Philosophical Studies,* 172(9), 2531–2549.

Engel, G. (1977). The need for a new medical model: A challenge for biomedicine. *Science*, 19(6), 129–136.

Esfeld, M. (2004). Quantum entanglement and a metaphysics of relations. *Studies in History and Philosophy of Modern Physics*, 35, 601–617.

Esfeld, M. & Lam, V. (2008). Moderate Structural Realism about Spacetime. *Synthese,* 160, 27–46.

Esfeld, M. & Lam, V. (2011). Ontic structural realism as a metaphysics of objects. In: A. and P. Bokulich (Eds.), *Scientific structuralism* (pp. 143–159). Dordrecht: Springer.

Esfeld, M. & Deckert, D. A. (2018). *A Minimalist Ontology of the Natural World*. New York: Routledge.

Farrell, J. & McClelland, T. (2017). Editorial: Consciousness and inner awareness. *Review of Philosophy and Psychology,* 8, 1–22.

Fasching, W. (2009). The Mineness of Experience. *Continental Philosophy Review*, 42(2), 131–48.

Fasching, W. (2011). "I Am of the Nature of Seeing": Phenomenological Reflections on the Indian Notion of Witness-Consciousness. In: M. Siderits, E. Thompson & D. Zahavi (Eds.), *Self, No Self? Perspectives from Analytical, Phenomenological, and Indian Traditions* (pp. 193–216). Oxford: Oxford University Press.

Fazakas, I., Bois, M. & Gozé, T. (2023). Giving thickness to the minimal self: coenesthetic depth and the materiality of consciousness. *Phenomenology and the Cognitive Sciences,* 1, 1–21.

Fleming, I. S. M., Frith, C., Goodale, M., Lau, H., LeDoux, J. E., et al. (2023, September 16). The Integrated Information Theory of Consciousness as Pseudoscience. *PsyArXiv*, 1, 1–9. https://doi.org/10.31234/osf.io/zsr78

Floridi, L. (2011). *The philosophy of information*. New York: Oxford University Press.

Fotopoulou, K. & Tsakiris, M. (2017). Mentalising homeostasis: the social origins of interoceptive inference. *Neuropsychoanalysis,* 19, 71–76.

Francescotti, R. (1999). How to define intrinsic properties. *Noûs*, 33(4), 590–609.

Frankish, K. (2016). Illusionism as a Theory of Consciousness, *Journal of Consciousness Studies*, 23(11–12), 11–39.

French, S. (2019). Defending Eliminative Structuralism and a Whole Lot More (or Less), *Studies in History and Philosophy of Science Part A*, 74, 22–29.

French, S. (2014). *The Structure of the World*. Oxford: Oxford University Press.

Frijda, N. H. (2007). *The laws of emotion*. Mahwah, NJ: Lawrence Erlbaum Associates, Inc.

Friston, K. J. (2010). The Free-Energy Principle: A Unified Brain Theory? *Nature Reviews Neuroscience*, 11(2), 127–138.

Friston, K. & Stephan, K. E. (2007). Free-energy and the brain. *Synthese*, 159(3), 417–458.

Friston, K. (2011). Embodied inference: Or "I think therefore I am, if I am what I think." In: W. Tschacher & C. Bergomi (Eds.), *The implications of embodiment (cognition and communication)* (pp. 89–125). Exeter: Imprint Academic.

Friston, K. (2013). Life as we know it. *Journal of the Royal Society Interface*, 10, 1–12.

Friston, K. (2018). Am I self-conscious? (or does self-organization entail self-consciousness?). *Frontiers in Psychology*, 9(579), 1–10.

Froese, T. & Di Paolo, E. A. (2009). Sociality and the life–mind continuity thesis. *Phenomenology and the Cognitive Sciences,* 8(4), 439–463.

Fuchs, T. & De Jaegher, H. (2009). Enactive intersubjectivity: Participatory sense-making and mutual incorporation. *Phenomenology and the Cognitive Sciences,* 8, 465–486.

Fuchs, T. (2018). *Ecology of the brain*. Oxford: Oxford University Press.

Gallagher, S. (2000). Philosophical conceptions of the self: Implications for cognitive science, *Trends in Cognitive Sciences,* 4(1), 14–21.

Gallagher, S. (2013). A pattern theory of self. *Frontiers in Human Neuroscience,* 7(443), 1–7.

Gallagher, S. (2021). Coherence in the self-pattern. In J. Noller (Ed.), *The unity of a person: Ontology — subjectivity — intersubjectivity* (pp. 127–145). London: Routledge.

Gallagher, S. & Zahavi, D. (2008). *The Phenomenological Mind*. London: Routledge.

Gallagher, S. (2022). Integration and Causality in Enactive Approaches to Psychiatry. *Frontiers in psychiatry*, 13, 1–16.

Gallagher, S., Raffone, A., Berkovich-Ohana, A. et al. (2023). The Self-Pattern and Buddhist Psychology. *Mindfulness*, 1, 1–10.

Gasser, G. & Stefan, M. (2012). *Personal Identity. Complex or Simple?* Cambridge: Cambridge University Press.

Gertler, B. (2012). Renewed acquaintance. In: D. Smithies & D. Stoljar (Eds.), *Introspection and consciousness* (pp. 89–123). Oxford: Oxford University Press.

Gibson, J. J. (1979). *The ecological approach to visual perception*. Boston: Houghton Mifflin.

Gipps, R. (2022). *On madness: Understanding the psychotic mind*. London: Bloomsbury Academic.

Godfrey-Smith, P. (2016). *Other Minds: The Octopus, the Sea, and the Deep Origins of Consciousness*. New York: Farrar, Straus and Giroux.

Goff, P. (2009). Why Panpsychism Doesn't Help Us Explain Consciousness. *Dialectica*, 63, 289–311.

Goff, P. (2011). A posteriori physicalists get our phenomenal concepts wrong. *Australasian Journal of Philosophy,* 89(2), 191–209.

Goff, P. (2015). Real acquaintance and physicalism. In: P. Coates & S. Coleman (Eds.), *Phenomenal Qualities: Sense, Perception, and Consciousness* (pp. 121–141). Oxford: Oxford University Press.

Goff, P. (2015). Against constitutive forms of Russellian monism. In: T. Alter & Y. Nagasawa (Eds.), *Consciousness in the Physical World: Perspectives on Russellian Monism* (pp. 370–401). Oxford: Oxford University Press.

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

Goff, P. (2021). Putting Consciousness First: Replies to Critics. *Journal of Consciousness Studies*, 28(9–10), 289–328.

Guillot, M. (2016). Thinking of oneself as the thinker: the concept of self and the phenomenology of intellection. *Philosophical Explorations*, 19(2), 138–160.

Guillot, M. (2017). I Me Mine: on a confusion concerning the subjective character of experience. *Review of Philosophy and Psychology: Special Issue on Consciousness and Inner Awareness,* 8(1), 23–53.

Gurwitsch, A. (1941). A Non-Egological Conception of Consciousness. *Philosophy and Phenomenological Research*, 1(3), 325–38.

Heft, H. (2001). *Ecological psychology in context: James Gibson, Roger Barker, and the legacy of William James's radical empiricism*. Hillsdale, NJ: Erlbaum.

Heinämaa, S., Lähteenmäki, V. & Remes, P. (2007). *Consciousness: From Perception to Reflection in the History of Philosophy*. Dordrecht: Springer.

Heinämaa, S. (2015). Anonymity and Personhood: Merleau-Ponty's account of the subject of perception. *Continental Philosophy Review,* 48, 123–142.

Henrich, D. (1970). Selbstbewußtsein, kritische Einleitung in eine Theorie. In: R. Bubner, K. Cramer & R. Wiehl (Eds.), *Hermeneutik und Dialektik* (pp. 257–284). Tübingen: Mohr.

Henrich, D. (1971/2007). Selbstsein und Bewußtsein. unpublished lecture, published in the e-journal *Philosophie der Psychologie*, 1, 1–19. http://www.jp.philo.at/texte/HenrichD1.pdf (accessed Jun 1, 2017).

Heras-Escribano, M. (2021). Pragmatism, enactivism, and ecological psychology: towards a unified approach to post-cognitivism. *Synthese,* 198(1), 337–363.

Hibbert, B. J. (2022). From quarks to cognition: in pursuit of a complete micropsychist theory of consciousness [Doctoral dissertation]. Keel: Keele University.

Hoffman, G. A. (2016). Out of our skulls: How the extended mind thesis can extend psychiatry. *Philosophical Psychology,* 29(8), 1160–1174.

Hohwy, J. (2013). *The predictive mind*. Oxford: Oxford University Press.

Hohwy, J. (2016). The self-evidencing brain. *Noûs,* 50(2), 259–85.

Hohwy, J. (2022). Conscious Self-Evidencing. *Review of Philosophy and Psychology*, 13, 809–828.

Hohwy, J. & Michael, J. (2017). Why would any body have a self? In: F. Vignemont & A. Alsmith (Eds.), *The Subject's matter: Self-consciousness and the body* (pp. 363–392). Cambridge, Mass.: MIT Press.

Horgan, T. & Kriegel, U. (2007). Phenomenal Epistemology: What Is Consciousness That We May Know It So Well? *Philosophical Issues,* 17(1), 123–144.

Horgan, T. & Nichols, S. (2016). The zero point and I. In: S. Miguens, G. Preyer & C. B. Morando (Eds.), *Pre-Reflective Consciousness: Sartre and Contemporary Philosophy of Mind* (pp. 143–176). London: Routledge.

Howell, R. & Thompson, B. (2017). Phenomenally mine: in search of the subjective character of consciousness. *Review of Philosophy and Psychology,* 8(1), 103–127.

Hume, D. (1978 /1740). *A treatise of human nature.* Oxford: Oxford University Press.

Humphreys, P. (1997). How properties emerge. *Philosophy of Science,* 64, 1–17.

Hutto, D. D. (2010). Radical enactivism and narrative practice: Implications for psychopathology. In: T. Fuchs, P. Henningsen & H. Sattel (Eds.),

Coherence and disorders of the embodied self (pp. 43–66). Stuttgart: Schattauer.

Hutto, D. D. & Myin, E. (2013). *Radicalizing enactivism: Basic minds without content*. Cambridge: The MIT Press.

James, W. (1890/1981). *The Principles of Psychology*. New York: Dover.

James, W. (1912). A World of Pure Experience. In his *Essays in Radical Empiricism* (pp. 39–91). New York: Longman Green and Co.

Jaspers, K. (1959/1963). *General Psychopathology*. Manchester: Manchester University Press.

Jerotić, S. & Nešić, J. (2023). Phenomenology and psychiatry: Shaping the diagnosis. *Theoria,* 66(1), 89–107.

Jonas, H. (1966/2001). *The phenomenon of life: Toward a philosophical biology*. Evanston, IL: Northwestern University Press.

Jonas, H. (1992). The burden and blessing of mortality. *The Hastings Center Report*, 22(1), 34–40.

Kim, J. (1982). Psychophysical Supervenience. *Philosophical Studies*, 41(1), 51–70.

Kiverstein, J. & Rietveld, E. (2018). Reconceiving representation-hungry cognition: an ecological-enactive proposal. *Adaptive Behavior,* 26, 147–163.

Kiverstein, J. (2020). Free energy and the self: an ecological–enactive interpretation. *Topoi*, 39, 559–574.

Klawonn, E. (1987). The "I": On the Ontology of First Personal Identity, *Danish Yearbook of Philosophy*, 24, 43–76.

Klin, A., Jones, W., Schultz, R. & Volkmar, F. (2003). The enactive mind, or from actions to cognition: Lessons from autism. *Philosophical Transactions of the Royal Society of London. Series B: Biological Sciences*, 358(1430), 345–360.

Kolchinsky, A. & Wolpert, D. H. (2018). Semantic information, autonomous agency and non-equilibrium statistical physics. *Interface Focus,* 8(6), 1–17.

Kriegel, U. (2009). *Subjective consciousness: A self-representational theory*. Oxford: Oxford University Press.

Krueger, J. (2020). Schizophrenia and the scaffolded self. *Topoi,* 39, 597–609.

Kyselo, M. (2014). The Body Social: An Enactive Approach to the Self. *Frontiers in Cognitive Science,* 5, 1–17.

Kyselo, M. (2016). The enactive approach and disorders of the self. The case of schizophrenia. *Phenomenology and the Cognitive Sciences*, 15(4), 591–616.

Kyselo, M. (2016). The minimal self needs a social update. *Philosophical Psychology*, 1, 1–9.

Kyselo, M. (2023). What Self in Self-Organization? Engaging Varela's Epistemology for the Co-Embodied Self. *Journal of Consciousness Studies,* 30(11), 80–103.

Ladyman, J. (1998). What is structural realism? *Studies in History and Philosophy of Science Part A*, 29(3), 409–424.

Ladyman, J., Ross, D., Spurrett, D. & Collier, J. (2007). *Every Thing Must Go: Metaphysics Naturalized*. Oxford: Oxford University Press.

Ladyman, J. (2020, Winter Edition). Structural Realism. In: E. N. Zalta (Ed.), *Stanford Encyclopedia of Philosophy*. https://plato.stanford.edu/archives/win2020/entries/structural-realism (accessed June 19, 2021).

Letheby, C. & Gerrans, P. (2017). Self unbound: ego dissolution in psychedelic experience. *Neuroscience of Consciousness*, 3(1), 1–11.

Levine, J. (1983). Materialism and Qualia: The Explanatory Gap. *Pacific Philosophical Quarterly,* 64(9), 354–361.

Levine, J. (2001). *Purple haze: The puzzle of consciousness*. Oxford: Oxford University Press.

Lewis, D. (1983). New Work for a Theory of Universals. *Australasian Journal of Philosophy,* 61, 343–377.

Lewis, D. (1986). *On the Plurality of Worlds.* New York: Blackwell.

Lewis, D. & Langton, R. (1998). Defining 'intrinsic'. *Philosophy and Phenomenological Research,* 58(2), 333–345.

Limanowski, J. & Blankenburg, F. (2013). Minimal self-models and the free energy principle. *Frontiers in Human Neuroscience*, 7, 1–12.

Lowe, E. J. (1996). *Subjects of Experience*. Cambridge: Cambridge University Press.

Lowe, E. J. (1998). *The Possibility of Metaphysics: Substance, Identity, and Time.* Oxford: Clarendon Press.

Lowe, E. J. (2003). Individuation. In: M. J. Loux & D. W. Zimmerman (Eds.), *The Oxford Handbook of Metaphysics* (pp. 75–95). Oxford: Oxford University Press.

Lowe, E. J. (2006). *The four-category ontology: A metaphysical foundation for natural science*. Oxford: Oxford University Press.

Lowe, E. J. (2012). The Probable Simplicity of Personal Identity. In: G. Gasser & M. Stefan (Eds.), *Personal Identity: Complex or Simple* (pp. 137–155). Cambridge: Cambridge University Press.

Mach, E. (1886/1959). *Die Analyse der Empfindungen und das Verhältnis des Physischen zum Psychischen*, *fifth edition translated as The Analysis of Sensations and the Relation of Physical to the Psychical*. New York: Dover.

Malafouris, L. (2013). *How things shape the mind*. Cambridge, MA: MIT Press.

Maiese, M. (2016). *Embodied selves and divided minds*. Oxford: Oxford University Press.

Maturana, H. & Varela, F. J. (1972/1980). *Autopoiesis and cognition: The realization of the living*. Dordrecht: D. Reidel.

Maturana, H. R. & Varela, F. J. (1987). *The Tree of Knowledge: The Biological Roots of Human Understanding*. Boston, MA: Shambhala Publications.

McGinn, C. (1989). Can We Solve the Mind-body Problem? *Mind,* 98, 349–66.

McQueen, K. J. (2019). Interpretation-Neutral Integrated Information Theory. *Journal of Consciousness Studies,* 26, 76–106.

Merleau-Ponty, M. (1945/1993). *Phénoménologie de la perception*. Paris: Gallimard.

Merleau-Ponty, M. (1945/2002). *Phenomenology of perception*. London: Routledge.

Merleau-Ponty, M. (1968/2003). *Nature: course notes from the College de France*. Evanston, IL: Northwestern University Press.

Merleau-Ponty, M. (1963). *The Structure of Behavior*. Pittsburgh, PA: Duquesne University Press.

Metzinger, T. (2003). *Being No One*. Cambridge: MIT Press.

Metzinger, T. (2011). The no-self alternative. In S. Gallagher (Ed.), *The Oxford Handbook of the Self* (pp. 279–296). Oxford: Oxford University Press.

Mitchell, S. W. (1866, July). The Case of George Dedlow. *The Atlantic Monthly*. http://www.gutenberg.org/etext/693 (accessed April 2, 2009).

Miguens, S., Preyer, G. & Morando, C. B. (2016). *Pre-Reflective Consciousness: Sartre and Contemporary Philosophy of Mind*. New York: Routledge.

Mindt, G. (2017). The Problem with the "Information" in Integrated Information Theory. *Journal of Consciousness Studies*, 24, 130–154.

Mindt, G. (2019). Reducing Uncertainty: Understanding the Information-Theoretic Origins of Consciousness [Doctoral dissertation]. Budapest: Central European University.

Mindt, G. (2021). Not All Structure and Dynamics Are Equal. *Entropy*, 23(9), 1226–1240.

Mørch, H. H. (2014). Panpsychism and Causation: A New Argument and a Solution to the Combination Problem [Doctoral dissertation]. Oslo: University of Oslo.

Mørch, H. H. (2018). Does Dispositionalism Entail Panpsychism? *Topoi,* 39, 1073–1088.

Molnar, G. (2003). *Powers: a study in metaphysics*. Oxford: Oxford University Press.

Nagel, T. (1974). What Is It Like to Be a Bat? *Philosophical Review*, 83(4), 435–450.

Nagel, T. (1979). Panpsychism. In his *Mortal questions* (pp. 181–196). Cambridge: Cambridge University Press.

Neisser, U. (1988). Five kinds of self-knowledge. *Philosophical Psychology*, 1(1), 35–59.

Nelson, B., McGorry, P. D. & Fernandez, A. V. (2021). Integrating clinical staging and phenomenological psychopathology to add depth, nuance, and utility to clinical phenotyping: a heuristic challenge. *The Lancet Psychiatry*, 8(2), 162–168.

Nešić, J. (2013a). Dispositional explanations in dualism. *Filozofija i društvo*, 24(4), 218–241.

Nešić, J. (2013b). Dualism, mental explanations and explanatory exclusion. *Theoria*, 56(1), 19–45.

Nešić, J. (2015). Problem mentalne kauzalnosti: rešenje preko dispozicione teorije kauzalnosti [Doktorska disertacija]. Beograd: Univerzitet u Beogradu.

Nešić, J. (2017). Against deflation of the subject. *Philosophy and Society*, 28(4), 1102–1121.

Nešić, J. (2018a). Does Integrated Information lack subjectivity? *Theoria*, 61(2), 131–145.

Nešić, J. (2018b). Subjectivity emergence in panpsychism. *Facta Universitatis, Series: Philosophy, Sociology, Psychology and History*, 17(3), 169–178.

Nešić, J. (2022a). Towards a Neutral-Structuralist Theory of Consciousness and Selfhood. *International Studies in the Philosophy of Science*, 35(3–4), 243–259.

Nešić, J. (2022b). Enactive framework for psychiatric disorders. *Engrami*, 44(1), 40–52.

Nešić, J. (2023a). Ecological-enactive account of autism spectrum disorder. *Synthese,* 201(67), 1–22.

Nešić, J. (2023b). I am mine: From phenomenology of self-awareness to metaphysics of selfhood. *Belgrade Philosophical Annual*, 36(1), 67–85.

Nešić, J. (2023c). Affording autistic persons epistemic justice. In: N. Cekić (Ed.), *Virtues and vices — between ethics and epistemology* (pp. 343–361). Belgrade: Faculty of Philosophy, University of Belgrade.

Newen, A., de Bruin, L. & S. Gallagher (Eds.) (2018). *Oxford handbook of cognition: Embodied, enactive, embedded and extended.* Oxford: Oxford University Press.

Ney, A. (2015). A Physicalist Critique of Russellian Monism. In A. Alter, & Y. Nagasawa (Eds.), *Consciousness in the Physical World* (pp. 346–369). Oxford: Oxford University Press.

Nida-Rümelin, M. (2007). Grasping phenomenal properties. In: T. Alter & S. Walter (Eds.), *Phenomenal Concepts and Phenomenal Knowledge: New Essays on Consciousness and Physicalism* (pp. 307–388). Oxford: Oxford University Press.

Nida-Rümelin, M. (2012). The non-descriptive individual nature of conscious beings. In: G. Gasser & M. Stefan (Eds.), *Personal Identity: Complex or Simple* (pp. 157–176). Cambridge: Cambridge University Press.

Nida-Rümelin, M. (2014). Basic intentionality, primitive awareness, and awareness of oneself. In: A. Reboul (Ed.), *Mind, Values and Metaphysics. Philosophical Papers dedicated to Kevin Mulligan — Volume 2* (pp. 261–290). London: Springer.

Nida-Rümelin, M. (2016). The experience property framework — a misleading paradigm. *Trends in philosophy of language and mind, special issue of Synthese*, 195(8), 3361–3387.

Nida-Rümelin, M. (2017). Self-Awareness. *Review of Philosophy and Psychology: Special Issue on Consciousness and Inner Awareness,* 8(1), 55–82.

Niiniluoto, I. (2019, Winter Edition). Scientific Progress. In E. N. Zalta (Ed.), *Stanford Encyclopedia of Philosophy*. https://plato.stanford.edu/entries/scientific-progress/ (accessed December 4, 2020).

Noë, A. (2004). *Action in Perception*. Cambridge, MA: MIT Press.

Northoff, G. (2018). *The Spontaneous Brain: From the Mind-Body to the World-Brain Problem*. Cambridge, MA: MIT Press.

O'Conaill, D. (2020). The identity of experiences and the identity of the subject. *Philosophical Studies,* 177, 987–1005.

O'Connor, T. (1994). Emergent properties. *American Philosophical Quarterly,* 31, 91–104.

Oizumi, M., Albantakis, L. & Tononi, G. (2014). From the Phenomenology of the Mechanisms of Consciousness: Integrated Information Theory 3.0. *PLoS Computational Biology*, 10(5), 1–25.

Olson, E.T. (1998). There is no problem of the self. *Journal of Consciousness Studies,* 5, 645–657.

Pace, Giannotta A. (2021a). Panqualityism as a critical metaphysics for neurophenomenology. *Constructivist Foundations,* 16(2), 163–166.

Pace, Giannotta A. (2021b). Autopoietic Enactivism, Phenomenology, and the Problem of Naturalism: A Neutral Monist Proposal. *Husserl Studies,* 37, 209–228.

Parfit, D. (1984). *Reasons and Persons*. Oxford: Oxford University Press.

Pereboom, D. (2015). Consciousness, Physicalism, and Absolutely Intrinsic Properties. In: T. Alter & Y. Nagasawa (Eds.), *Consciousness in the Physical World: Perspectives on Russellian Monism* (pp. 300–324). Oxford: Oxford University Press.

Perrykkad, K. & Hohwy, J. (2020). Modelling Me, Modelling You: the Autistic Self. *Review Journal of Autism and Developmental Disorders,* 7, 1–31.

Pothast, U. (1971). *Über einige Fragen der Selbstbeziehung*. Frankfurt am Main: Vittorio Klostermann.

Powys, J. C. (1930). *In Defence of Sensuality*. London: Victor Gollancz Ltd.

Powys, J. C. (1930). *The Meaning of Culture*. London: Jonathan Cape Ltd.

Priest, S. (2000). *The Subject in Question: Sartre's Critique of Husserl in 'The Transcendence of the Ego*'. London: Routledge.

Rietveld, E. (2008). *Unreflective action: a philosophical contribution to integrative neuroscience*. Amsterdam: ILLC- Dissertation Series.

Roelofs, L. (2015). Combining Minds: A Defence of the Possibility of Experiential Combination [Doctoral dissertation]. Toronto: University of Toronto.

Roberts, T., Krueger, J. & Glackin, S. (2019). Psychiatry beyond the brain: Externalism, mental health, and autistic spectrum disorder. *Philosophy, Psychiatry and Psychology*, 26(3), 51–68.

Rosenberg, G. (1998). The Boundary Problem for Phenomenal Individuals. In: S. Hameroff, A. Kaszniak & A. Scott (Eds.), *Toward a Science of Consciousness: The First Tucson Discussions and Debates (Complex Adaptive Systems)* (pp. 1–23). Tucson: MIT Press.

Rosenberg, G. (2004). *A Place for Consciousness: Probing the Deep Structure of the Natural World*. Oxford: Oxford University Press.

Russell, B. (1912). *The Problems of Philosophy*. Oxford: Home University Library.

Russell, B. (1921). *The Analysis of Mind*. London: Routledge.

Russell, B. (1927). *The Analysis of Matter*. Nottingham: Spokesman.

Sacks, O. (1987). *The Man Who Mistook His Wife for a Hat and Other Clinical Tales*. New York: Perennial Library.

Saks, Elyn R. (2007). *The Centre Cannot Hold (My Journey Through Madness)*. New York: Hyperion.

Sass, L., Borda, J.P., Madeira, L., Pienkos, E. & Nelson, B. (2018). Varieties of self disorder: a bio-pheno-social model of schizophrenia. *Schizophrenia Bulletin,* 44(4), 720–727.

Sayre, K. (1976). *Cybernetics and the Philosophy of Mind*. Atlantic Highlands: Humanities Press.

Schilbach, L., Timmermans, B., Reddy, V., Costall, A., Bente, G., Schlicht, T. & Vogeley, K. (2013). Toward a second-person neuroscience. *Behavioral and Brain Sciences*, 36(4), 393–414.

Schneider, K. (1959). *Clinical Psychopathology*. New York: Grune & Stratton.

Schwitzgebel, E. (2014). Why Tononi Should Think That the United States Is Conscious. http://schwitzsplinters.blogspot.com/2012/03/why-tononi-should-think-that-united (accessed June 15, 2016).

Seager, W. (2006). The 'intrinsic nature' argument for panpsychism. *Journal of Consciousness Studies,* 13(10–11), 129–145.

Seager, W. (2010). Panpsychism, Aggregation and Combinatorial Infusion. *Mind and Matter*, 8(2), 167–184.

Seager, W. (2019). *The Routledge Handbook of Panpsychism*. New York: Routledge.

Seth., A. (2013). Interoceptive inference, emotion and the embodied self. *Trends in Cognitive Sciences,* 17(11), 565–573.

Seth, A. (2021). *Being you: A new science of consciousness*. London: Faber & Faber.

Siderits, M. Thompson, E. & Zahavi, D. (2011). *Self, no self?: perspectives from analytical, phenomenological, and Indian traditions*. Oxford: Oxford University Press.

Siewert, C. (2013). Phenomenality and self-awareness. In: U. Kriegel (Ed.), *Phenomenal Intentionality* (pp. 235–258). Oxford: Oxford University Press.

Searle, J. F. (2013, January 10). Can information theory explain consciousness? [Review of the book *Consciousness: confessions of a romantic reductionist,* by Kristof Koch]. *The New York Review of Books*. https://www.nybooks.com/articles/2013/01/10/can-information-theory-explain-consciousness/ (accessed May 2, 2015).

Shanahan, M. (2015). Ascribing Consciousness to Artificial Intelligence, *arXiv,* 1504.05696, 1–16. https://doi.org/10.48550/arXiv.1504.05696.

Shani, I. (2015). Cosmopsychism: A Holistic Approach to the Metaphysics of Experience. *Philosophical Papers,* 44(3), 389–437.

Shannon, C. E. (1948). A Mathematical Theory of Communication. *The Bell System Technical Journal,* 27, 379–423.

Shoemaker, S. (2002). Kim on Emergence. *Philosophical Studies,* 58, 53–63.

Shoemaker, S. (1996). *The first-person perspective and other essays*. Cambridge: Cambridge University Press.

Shrader, W. (2010). Shoemaker on Emergence. *Philosophical Studies,* 150, 285–300.

Slors, M. V. P. & Jongepier, F. (2014). Mineness without Minimal Selves. *Journal of Consciousness Studies,* 21(7–8), 193–219.

Stace, W. T. (1961). *Mysticism and philosophy*. New York: St. Martin's Press.

Stoljar, D. (2017). Physicalism. In E. N. Zalta (Ed.), *Stanford Encyclopedia of Philosophy*. https://plato.stanford.edu/archives/sum2021/entries/physicalism/ (accessed January 23, 2021).

Strawson, G. (2009). *Selves: An Essay in Revisionary Metaphysics.* Oxford: Oxford University Press.

Swinburne, R. (1995). Thisness. *Australasian Journal of Philosophy*. 73(3), 389–400.

Strawson, P. F. (1959). *Individuals: An Essay in Descriptive Metaphysics*. London: Methuen.

Strawson, G. (2008). *Real Materialism and Other Essays*. Oxford: Oxford University Press.

Stubenberg, L. (2016). Neutral Monism. In E. N. Zalta (Ed.), *Stanford Encyclopedia of Philosophy*. https://plato.stanford.edu/archives/fall2018/entries/neutral-monism/ (accessed February 22, 2021).

Tammet, D. (2006). *Born on a Blue Day: A Memoir of Asperger's and an Extraordinary Mind*. London: Hodder and Stoughton.

Thompson, E. & F. J. Varela (2001). Radical embodiment: Neural dynamics and consciousness. *Trends in Cognitive Sciences,* 5(10), 418–25.

Thompson, E. (2007). *Mind in Life: Biology, Phenomenology, and the Sciences of Mind*. Cambridge, MA: Harvard University Press.

Thompson, E. (2015). *Waking, Dreaming, Being: Self and Consciousness in Neuroscience, Meditation, and Philosophy*. New York: Columbia University Press.

Thompson, E. (2020). *Why I am not a Buddhist*. New Haven and London: Yale University Press.

Tononi, G. (2008). Consciousness as Integrated Information: a Provisional Manifesto. *The Biological Bulletin*, 215, 216–242.

Tononi, G. (2012). Integrated Information Theory of Consciousness: an Updated Account. *Archives Italiennes de Biologie*, 150, 290–326.

Tononi, G. & Koch, C. (2015). Consciousness: here, there and everywhere? *Philosophical Transactions of the Royal Society B,* 370(1668), 1–18.

Tononi, G., Boly, M., Massimini, M. & Koch, C. (2016). Integrated information theory: from consciousness to its physical substrate. *Nature Reviews Neuroscience,* 17, 450–461.

Turausky, K. (2014). Wherever You Go, There You Are: On Individuative Subjective Phenomenology. *Essays in Philosophy*, 15(2), 249–265.

Varela, F. J., Thompson & E., Rosch, E. (1991). *The embodied mind: Cognitive science and human experience* (6th ed.). Cambridge, MA: MIT Press.

Varela, F. J. (1991). Organism: A Meshwork of Selfless Selves. In: A. I. Tauber (Ed.), *Organism and the Origins of Self. Boston Studies in the Philosophy of Science* (pp. 79–109). Dordrecht: Springer.

Varela, F. J. (1996). Neurophenomenology: A methodological remedy for the hard problem. *Journal of Consciousness Studies,* 3(4), 330–49.

Varela, F. J. (1997). Patterns of life: Intertwining identity and cognition. *Brain and Cognition*, 34(1), 72–87.

von Uexküll, J. (1909). *Umwelt und Innenwelt der Tiere*. Berlin: Springer.

Walsh, P. J. (2020). Intercorporeity and the first-person plural in Merleau-Ponty. *Continental Philosophy Review,* 53, 21–47.

Williams, D. (1994). *Somebody Somewhere: Breaking Free from the World of Autism*. New York: Times Books.

Williford, K. (2015). Representationalisms, Subjective Character, and Self-Acquaintance. In T. Metzinger & J. M. Windt (Eds.), *Open MIND*, 39(T) (pp. 1–27). Frankfurt am Main: MIND Group.

Worrall, J. (1989). Structural realism: the best of both worlds? *Dialectica,* 43(1–2), 99–124.

Yergeau, M. (2012). "I stim, therefore I am" (video). *YouTube*, uploaded January 26, 2012. https://www.youtube.com/watch?v=s2QSvPIDXwA (accessed June 4, 2023).

Zahavi, D. (1999). *Self-awareness and alterity: A phenomenological investigation*. Evanston, Illinois: Northwestern University Press.

Zahavi, D. (2005). *Subjectivity and Selfhood: Investigating the First-Person Perspective*. Cambridge: MIT Press.

Zahavi, D. (2014). *Self and Other: Exploring Subjectivity, Empathy, and Shame*. Oxford: Oxford University Press.

Zahavi, D. & Kriegel, K. (2015). For-Me-Ness: What It is and What It is Not. In: D. Dahlstrom, A. Elpidorou & W. Hopp (Eds.), *Philosophy of Mind and Phenomenology* (pp. 36–53). New York and London: Routledge.

CIP – Каталогизација у публикацији
Народна библиотека Србије, Београд

165.12
165.242.1:165.62

NEŠIĆ, Janko, 1985–
Self : from one to many and back to none / Janko Nešić. – Belgrade : Institute of Social Sciences, 2024 (Beograd : Faculty of Technology and Metallurgy, Research and Development Centre of Printing Technology). – 148 str. ; 21 cm. – (Series Monographs / [Institute of Social Sciences, Belgrade])

Tiraž 150. – Napomene i bibliografske reference uz tekst. – Bibliografija: str. 133–148.

ISBN 978-86-7093-274-6

а) Самосвест б) Свест в) Феноменологија г) Субјективност
COBISS.SR-ID 143396873

Language proofreading: M. Prevodi | **Graphic design and layout:** MITRACK | **Print:** 150 copies | **Printed by:** Faculty of Technology and Metallurgy, Research and Development Centre of Printing Technology